AD_____

AWARD NUMBER DAMD17-98-1-8194

TITLE: Computerized Analysis of MR and Ultrasound Images of Breast Lesions

PRINCIPAL INVESTIGATOR: Maryellen L. Giger, Ph.D.

CONTRACTING ORGANIZATION: The University of Chicago
Chicago, Illinois 60637

REPORT DATE: July 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1999	Annual (1 Jul 98 - 30 Jun 99)

4. TITLE AND SUBTITLE
Computerized Analysis of MR and Ultrasound Images of Breast Lesions

5. FUNDING NUMBERS
DAMD17-98-1-8194

6. AUTHOR(S)
Maryellen L. Giger, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of Chicago
Chicago, Illinois 60637

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We are developing methods for the classification of lesions in ultrasound and MR images of the breast. We are retrospectively collecting large datasets of ultrasound and MR images of the breast with solid pathology and radiologists' ratings. These cases include malignant lesions, benign lesions, and complex cysts. We are developing noninvasive computerized methods for characterizing the lesions to yield an output related to the probability of malignancy. We plan to evaluate the efficaciesof the new image analysis methods inthe task of distinguishing between malignant and benign lesions. It is expected that the results from this research will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer computer-aided diagnosis, MR image, ultrasound	53 16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

__✓__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Maryellen Sage 8/5/89

PI - Signature Date

Table of Contents

INTRODUCTION

The main hypothesis to be tested is that, computerized analysis of breast ultrasound and MR images should yield new methods for distinguishing between malignant and benign lesions and thus, reduce the number of unnecessary biopsies. In addition, even higher performance is expected when a combination of features from mammographic, MR, and ultrasound images is used as an aid to radiologists in the task of distinguishing between malignant and benign lesions. The main goal of the proposed research is to develop noninvasive, computerized methods for analyzing ultrasound and MR (magnetic resonance) images of breast lesions to aid radiologists in their workup of suspect lesions. The specific objectives of the research to be addressed are: 1. Create a database of ultrasound and MR images including malignant lesions, benign solid masses, and complex cysts; 2. Develop noninvasive, computerized methods for characterizing the lesions to yield an output related to the probability of malignancy; and 3. Evaluate the efficacies of the new image analysis methods in the task of distinguishing between malignant and benign lesions. It is expected that the results from this research will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy. Computerized image analysis techniques that can objectively and reliably classify lesions based upon reported sonographic and/or MR characteristics of benign and malignant masses, especially if combined with their mammographic features, could significantly improve the specificity of breast imaging and the evaluation of breast masses. The proposed work is novel in that computer-aided diagnosis techniques applied to gray-scale sonographic images has not yet been reported. In addition, computerized analysis of MR images of the breast has mainly been limited to only temporal analysis using contrast media.

BODY

Breast cancer is a leading cause of death in women, causing an estimated 46,000 deaths per year (1). Mammography is the most effective method for the early detection of breast cancer, and it has been shown that periodic screening of asymptomatic women does reduce mortality (2-4). Many breast cancers are detected and referred for surgical biopsy on the basis of a radiographically detected mass lesion or cluster of microcalcifications. Although general rules for the differentiation between benign and malignant mammographically identified breast lesions exist (5, 6), considerable misclassification of lesions occurs with the current methods. On average, less than 30% of masses referred for surgical breast biopsy are actually malignant (7).

Breast sonography is used as an important adjunct to diagnostic mammography and is typically performed to evaluate palpable and mammographically identified masses in order to determine their cystic vs. solid natures. The accuracy of ultrasound has been reported to be 96-100% in the diagnosis of simple benign cysts (8). Masses so characterized do not require further evaluation; however, 75% of masses prove to be indeterminate or solid on sonography and are candidates for further intervention (9). With the advent of modern high-frequency transducers that have improved spatial and contrast resolution, a number of sonographic features have emerged as potential indicators of malignancy, while other features are typical for benign masses (10,11). Benign features include hyperechogenicity, ellipsoid shape, mild lobulation, and a thin, echogenic pseudocapsule. Malignant features include spiculation, angular margins, marked hypoechogenicity, posterior acoustic shadowing, and a depth:width ratio greater than 0.8. Recently, Stavros, et al., used these and other features to characterize masses as benign, indeterminate, and malignant (12). Their classification scheme had a sensitivity of 98.4% and a negative predictive value of 99.5%. However, the sonographic evaluation described by these investigators is much more extensive and complex than is usually performed at most breast imaging centers.

Breast MR imaging as an adjunct to mammography and sonography reveals breast cancer with a higher sensitivity than do mammography and sonography only (13). However, using all three methods in the human interpretation process yielded a lower specificity. It also has been shown that temporal analysis from dynamic MR correlates with intensity of fibrosis in fibroadenomas (14).

Some computerized analyses of spatial features are being performed. Adams et al. achieved a separation between malignant and benign lesions using a statistical analysis, however, their database consisted of only 16 cases (15).

Computerized image analysis techniques that can objectively and reliably classify lesions based upon reported sonographic and/or MR characteristics of benign and malignant masses, especially if combined with their mammographic features, could significantly improve the specificity of breast imaging and the evaluation of breast masses. Computer-aided techniques have been applied to the color Doppler evaluation of breast masses with promising results (16). However, color Doppler imaging is a technique which focuses only upon the vascularity of lesions. Since not all sonographically visible cancers have demonstrable neovascularity, this technique is inherently somewhat limited. On the other hand, computer-aided diagnosis techniques applied to gray-scale sonographic images has not yet been reported. In addition, computerized analysis of MR images of the breast has mainly been limited to only temporal analysis using contrast media.

Comprehensive summaries of investigations in the field of mammography CAD have been published by the co-P.I. (17, 18). In the 1960s and 70s, several investigators attempted to analyze mammographic abnormalities with computers. These previous studies demonstrated the potential capability of using a computer in the detection of mammographic abnormalities. Gale et al. (19) and Getty et al. (20) are both developing computer-based classifiers, which take as input diagnostically-relevant features obtained from radiologists' readings of breast images. Getty et al. found that with the aid of the classifier, community radiologists performed as well as unaided expert mammographers in making benign-malignant decisions. Swett et al. (21) are developing an expert system to provide visual and cognitive feedback to the radiologist using a critiquing approach combined with an expert system. At the University of Chicago, we have shown that the computerized analysis of mass lesions (22) and clustered microcalcifications (23) on digitized mammograms yields performances similar to an expert mammographer and significantly better than average radiologists in the task of distinguishing between malignant and benign lesions.

The proposed work is novel in that computer-aided diagnosis techniques have not yet been applied to gray-scale breast ultrasound and/or MR images. In addition, future research involving the use computers to merge features from mammographic, MR, and ultrasound images, as an aid to radiologists, has not yet been investigated.

The main goal of the proposed research is to develop noninvasive, computerized methods for analyzing ultrasound and MR (magnetic resonance) images of breast lesions to aid radiologists in their workup of suspect lesions. The specific objectives of the research to be addressed are: 1. Create a database of ultrasound and MR images including malignant lesions, benign solid masses, and complex cysts; 2. Develop noninvasive, computerized methods for characterizing the lesions to yield an output related to the probability of malignancy; and 3. Evaluate the efficacies of the new image analysis methods in the task of distinguishing between malignant and benign lesions. It is expected that the results from this research will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy.

1. Establishment of a database of ultrasound and MR images

Methods

Approximately 500 sonographically demonstrated lesions will be collected which will include aspirated complex cysts, and biopsied solid benign and malignant masses. The database of these collected cases will include the MR, sonographic, and mammographic images as well as the lesions' ultimate dispositions and diagnoses. (Note that funding already exists for the computerized analysis of the mammographic lesions). Based upon our current case load, we estimate that approximately 30% of the lesions will be complex cysts which required aspiration to prove their cystic nature, 40% will be benign solid masses, and 30% will be cancers. Palpable and mammographically identified

masses are evaluated sonographically by representative images in orthogonal planes, obtaining measurements in these same planes, and most masses are also evaluated with color Doppler imaging. Although the preliminary studies on ultrasound images involved the digitization of ultrasound films, the ultrasound images in this new database will be obtained directly from an ATL ultrasound machine, which produces digital image data. In addition, approximately 50 cases of MR images of the breast will be collected with a T1-weighted sequence, using coronal slices. After injection of GD contrast, 4 to 6 scans of both breasts will be obtained at 90 sec. time intervals. Biospy results will be used to determine truth regarding malignancy.

Results to Date

We currently have retrospectively collected over 300 ultrasound cases of mass lesions, all that had gone on to either biopsy or cyst aspiration. The images are obtained from University of Chicago and Northwestern University. The images are transferred in digital format from the ATL unit. The digital images within the ATL unit are obtained by screen capture. For each case we have at least two views of the lesion. We are currently collecting the corresponding mammograms for the study.

We currently have retrospectively collected 35 coronal MR cases from University of Muenster and 362 saggital MR cases of the breast from University of Pennsylvania. These are all volume datasets. Of the 362 saggital cases, 253 are focal (192 malignant, 51 benign, 10 normal), 74 are diffuse lesions (48 malignant and 25 benign), 10 are ductal (9 malignant and 1 benign), and 25 showed no enhancement (3 malignant, 19 benign, 3 normal).

2. Development of computerized method for the classification of lesions

The computerized method will include the image analysis of the texture within the lesion, the analysis of the margin of the lesion, and a comparison of the lesion with its surrounding tissue. Computerized analysis of the texture pattern in the lesion will be based on various texture analysis methods we have been investigating in our laboratory including Fourier spectra analysis and artificial neural networks. We note that it is extremely important to understand the relationship between the mathematical texture measures and the physical nature of the breast parenchyma.

The computerized analysis of the margin characteristics (edge definition) will involve feature extraction using radial edge-gradient analysis. We have done similar analysis on radiographic masses in determining their margin characteristics (spiculated and ill defined) (22). Two promising measures are the FWHM and the average radial gradient which correspond to the degree of spiculation and how ill-defined is the margin, respectively. From the computer-extracted margin, we will also determine the shape and irregularity of each lesion.

Specifically for the ultrasound images, comparison of the "density" and the texture patterns of the lesion with neighboring regions, including those deep to the lesion, will be performed in order to quantify its echogenicity and the amount of any posterior acoustic shadowing or enhancement. This will be performed by comparing feature values "below" the lesion to those obtained along side and below the lesion.

Temporal features will be determined from analyzing the MR image data over time. The contrast meduim uptake curve will be analyzed at various spatial locations within and around the suspect lesion. Temporal operators include the maximum uptake, mean gradient of uptake, and rms variation. Both two-dimensional and three-dimensional features will be calculated, e.g., irregularity and margin gradient characteristics. In addition, the spatial features will be investigated as a function of time.

We plan to use artificial neural networks along with other measures of the mass in question to obtain an estimate of the likelihood of malignancy. We will investigate merging the ultrasound image features and MR features with those from mammographic images of the same lesion. We already have funding support for the investigation involving radiographic imaging of masses.

The various features will serve as input data and will be supplied to the input units of the artificial neural network. Prior to input to the ANN, the features will be normalized between 0 and 1. The output data from the neural network are then obtained through successive nonlinear calculations in the hidden and output layers. The calculation at each unit in a layer includes a weighted summation of all entry numbers, an addition of a certain offset number, and a conversion into a number ranging from 0 to 1 using a sigmoid-shape function such as a logistic function. The neural network will be trained by a back-propagation algorithm using pairs of training input data and desired output data. The desired output data will be initially 1 if features of a malignant lesion are input and 0 otherwise. Once trained, the neural network will accept features of a lesion and will output a value that will be related to a likelihood of malignancy. Feature selection will be performed by analyzing the average and standard deviation of the various features for both malignant and benign lesions. Az values will be calculated for each feature as well as for the output of the ANNs. In addition, genetic algorithms, which we have used, in a pilot study, for optimizing feature selection for the task of distinguishing true-positive and false-positive mass detections, will also be used.

Results to Date: Ultrasound

We are developing computerized analyses of breast lesions in ultrasound images to aid in the discrimination between malignant and benign lesions (24). We extracted and calculated features related to lesion margin, shape, homogeneity (texture) and the nature of the posterior acoustic attenuation pattern in ultrasound images of the breast. Our database contained 184 digitized ultrasound images from 58 patients with 78 lesions. Benign lesions were confirmed by biopsy, cyst aspiration, or image interpretation alone, while malignant lesions were confirmed by biopsy. ROC analysis was used to study the performance of the various individual features and the output from linear discriminant analysis in distinguishing benign from malignant lesions. From ROC analysis, the feature characterizing the margin yielded A_Z values of 0.85 and 0.75, in the task of distinguishing between benign and malignant lesions in the entire database and in an equivocal database, respectively. The "equivocal" database contained lesions that had been proven to be benign or malignant by either cyst aspiration or biopsy. Linear discriminant analysis round-robin runs yielded A_Z values of 0.94 and 0.87 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively.

We are currently evaluating the method on a database of ultrasound images from Northwestern University. The database of over 300 cases includes pathology truth as well as radiologists BI-RADS ratings with all cases having gone to biopsy or aspiration. Our previous method required radiologists' manually-drawn lesion contours as input to the computerized classification scheme. The current method, however, involves automatic segmentation of the lesion contour from the ultrasound image data. Similar classification performance levels (in terms of A_Z in the task of distinguishing between benign and malignant lesions) were found for the method using either the manually-drawn contours or the computer-drawn contours. Results indicate that computerized analysis of ultrasound images has the potential to increase the specificity of breast sonography.

Results to Date: MRI

We are developing computerized analyses of breast lesions in MR images to aid in the discrimination between malignant and benign lesions (25). Dynamic MR data was obtained from 27 patients by a T1-weighted sequence, using 64 coronal slices, a typical slice thickness of 2 mm, and a pixel size of 1.25 mm. After injection of GDTPA contrast, 4 to 6 scans of both breasts were obtained at 90 sec. time intervals. The database contained 13 benign and 15 malignant lesions. Our computerized classification method includes temporal features of normalized speed and inhomogeneity of uptake, and spatial features of margin descriptors such as circularity and irregularity. Our results indicate that classification based on temporal and spatial features combined can yield a positive predictive value of 94%, and has the potential to reduce the number of unnecessary biopsies by approximately 92%.

We have developed a new method for automatically extracting the lesion from the 3D image set of the breast. Our previous results were based on the use of manually-drawn lesion contours in the various slices of the MR data. The new segmentation method involves the use of an encompassing shell to limit the region for local thresholding. ROC analysis yielded A_Z values of 0.90 when the manual segmentation was used in the classification and 0.93 when automatic segmentation was included.

We are currently evaluating the method on 362 cases from the University of Pennsylvannia. These images differ from our initial database in that the new cases are saggital and had fat suppression applied.

3. Evaluation in the task of distinguishing between malignant and benign lesions.

In order to test the capability of the neural networks to learn the features of malignant and benign lesions, a consistency test will be conducted in which the network is first trained with all the cases in the database and then tested with the same cases used in the training. A consistency test indicates that the network is able to "remember" all of the input types that were used for training. However, it is more important to test if the network can learn a generalized set of inputs from the examples provided and if it can then make a correct prediction for new cases that were not included in the training. Thus, a round-robin method will be employed to test the network's generalizing ability. With the jack-knife method, all but one of the cases are selected randomly from the database for training of the network, and the remaining one case is used for testing the network. The output values are then compared to the "truth" data. Various combinations of training and testing pairs will be selected by using a random number generator and the results will be analyzed using ROC analysis. ROC curves will be obtained by fitting continuous output data from the neural networks using the LABROC4 program (26). The area under the ROC curve (Az) will be used as an indicator of performance. In order to determine the structure of the neural network as well as the necessary number of training iterations, we will analyze the consistency results and the round-robin results in terms of Az as a function of number of iterations, momentum, learning rate and number of hidden units. We use Az as an indicator of performance since it includes information on both the sensitivity and specificity of the measures.

The proposed techniques are expected to yield measures about the likelihood of malignancy. Receiver Operating Characteristic (ROC) analysis (26) will be employed in evaluating the performance of the measures. We have used ROC analysis successfully in both evaluting the performance of human observers as well as that of computerized schemes. The task in which the image features will be evaluated will be in their ability to determine an estimate of the likelihood of malignancy. The decision variable for the ROC analysis will be each individual feature as well as combined measures within a modality and combined measures from multiple modalities (x-ray, MR, and ultrasound).

We expect that 500 lesions and their ultrasound images will be available for testing. Note that here the measure of performance will be the A_Z value (from ROC analysis) obtained in the task of distinguishing between malignant and benign lesions. To obtain an estimate of the number of lesions needed for adequate statistical power in testing differences in A_Z values, we assume only a correlation of 0.60 between the estimates of A_Z that are found for our current method involving the computerized analysis of mammographic lesions (A_Z=0.87) and that for the expected improved method (A_Z=0.92). With N_{pos} patients who have a malignant lesion and N_{neg} patients who have a benign lesion, the standard error of the resulting estimate can be approximated (Eqn. 9 in Ref. 27) by the expression $\{[2A_Z-(1-f)(1-A_Z)](1-A_Z)/3N_{pos}]\}^{1/2}$, where f represents the ratio N_{pos}/N_{neg}. Thus, with f = 1, the statistical power at a critical significance level of a = 0.05 for 500 mass lesions is 94%.

Results to Date

The evaluation of the methods will be performed in later years as indicated in the SOW.

KEY RESEARCH ACCOMPLISHMENTS

- Development of robust features for characterizing lesions in ultrasound images of the breast.
- Development of robust features for characterizing lesions in MRI images of the breast.
- Investigation and development of methods for segmentation in 2D and 3D image sets.

REPORTABLE OUTCOMES

1. Gilhuijs KGA, Giger ML, Bick U: Automated analysis of breast lesions in three dimensions using dynamic magnetic resonance imaging. Medical Physics 25:1647-1654, 1998.

2. Giger ML, Al-Hallaq H, Huo A, Moran C, Wolverton DE, Chan CW, Zhong W: Computerized analysis of lesions in ultrasound images of the breast. Academic Radiology (in press).

CONCLUSIONS

We have made great strides in the development of methods for the claslsification of lesions in ultrasound and MR images of the breast. We are retrospectively collecting large datasets of ultrasound and MR cases with solid pathology truth and radiologists' ratings. These cases include malignant lesions, benign solid masses, and complex cysts. We are developing noninvasive, computerized methods for characterizing the lesions to yield an output related to the probability of malignancy and plan to evaluate the efficacies of the new image analysis methods in the task of distinguishing between malignant and benign lesions. It is expected that the results from this research will aid radiologists in determining the likelihood of malignancy and in reducing the number of benign cases sent to biopsy.

REFERENCES

1. American Cancer Society: Cancer Facts and Figures -- 1995. Atlanta: American Cancer Society, 1995.
2. Feig SA: Decreased breast cancer mortality through mammographic screening: Results of clinical trials. Radiology 167:659-665, 1988.
3. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O: Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 30:187-210, 1992.
4. Smart CR, Hendrick RE, Rutledge JH, Smith RA: Benefit of mammography screening in women ages 40 to 49 years: Current evidence from randomized controlled trials. Cancer 75:1619-26, 1995.
5. Bassett LW, Gold RH: Breast Cancer Detection: Mammography and Other Methods in Breast Imaging New York: Grune and Stratton, 1987.
6. Kopans DB: Breast Imaging. Philadelphia: JB Lippincott, 1989.
7. Brown ML, Houn F, Sickles EA, Kessler LG: Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. AJR 165:1373-1377, 1995.
8. Jackson VP: The role of US in breast imaging. Radiology 177:305-311, 1990.
9. Hilton SW, Leopold GR, Olson LK, Wilson SA: Real-time breast sonography: application in 300 consecutive patients. AJR 147:479-486, 1986.
10. Tohno E, Cosgrove DO, Sloane JP: Ultrasound Diagnosis of Breast Diseases. Churchill Livingstone, Edinburgh, 1994, pp. 50-73.

11. Fornage BD, Lorigan JG, Andry E: Fibroadenoma of the breast: sonographic appearance. Radiology 172:671-675, 1989.
12. Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA: Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 196:123-134, 1995.
13. Muller-Schimpfle M, Stoll P, Stern W. et al.: Do mammography, sonography, and MR mammography have a diagnostic benefit compared with mammography and sonography? AJR 168: 1323-1329, 1997.
14. Brinck U, Fischer U, Korabiowska M, et al.: The variability of fibroadenoma in contrast-enhanced dynamic MR mammography. AJR 168: 1331-1334, 1997.
15. Adams AH, Brookeman JR, Merickel MB: Breast lesion discrimination using statistical analysis and shape measures on magnetic resonance imagery. Comp Med Imaging and Graphics 15: 339-349, 1991.
16. Huber S, Delorme S, Knopp MV, Junkermann H, Zuna I, von Fournier D, van Kaick G: Breast tumors: computer-assisted quantitative assessment with color Doppler US. Radiology 192:797-801, 1994.
17. Giger ML: Computer-aided diagnosis. In: Syllabus: A Categorical Course on the Technical Aspects of Breast Imaging, edited by Haus A, Yaffe M. Oak Brook, IL: RSNA Publications, 1993, pp. 272-298.
18. Vyborny CJ, Giger ML: Computer vision and artificial intelligence in mammography. AJR 162:699-708, 1994.
19. Gale AG, Roebuck EJ, Riley P, Worthington BS, et al.: Computer aids to mammographic diagnosis. British Journal of Radiology 60: 887-891, 1987.
20. Getty DJ, Pickett RM, D'Orsi CJ, Swets JA: Enhanced interpretation of diagnostic images. Invest. Radiol. 23: 240-252, 1988.
21. Swett HA, Miller PA: ICON: A computer-based approach to differential diagnosis in radiology. Radiology 163: 555-558, 1987.
22. Huo Z, Giger ML, Vyborny CJ, Bick U, Lu P, Wolverton DE, Schmidt RA: Analysis of spiculation in the computerized classification of mammographic masses" Medical Physics 22:1569-1579, 1995.
23. Jiang Y, Nishikawa RM, Wolverton DE, Giger ML, Doi K, Schmidt RA, Vyborny CJ: Automated feature analysis and classification of malignant and benign clustered microcalcifications. Radiology 198(3):671-678, 1996.
24. Giger ML, Al-Hallaq H, Huo A, Moran C, Wolverton DE, Chan CW, Zhong W: Computerized analysis of lesions in ultrasound images of the breast. Academic Radiology (in press).
25. Gilhuijs KGA, Giger ML, Bick U: Automated analysis of breast lesions in three dimensions using dynamic magnetic resonance imaging. Medical Physics 25:1647-1654, 1998.
26. Metz CE: Some practical issues of experimental design and data analysis in radiologicalROC studies. Invest. Radiol. 24: 234-245, 1989.
27. Bamber D. The area above the ordinal dominance graph and the area below the receiver operating graph. J. Math Psych 12: 387-415, 1975.

Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging

Kenneth G. A. Gilhuijs[a] and Maryellen L. Giger
The Kurt Rossmann Laboratories for Radiologic Image Research, The University of Chicago, Department of Radiology, MC 2026, 5841 South Maryland Avenue, Chicago, Illinois 60637

Ulrich Bick[b]
The University of Muenster, Department of Radiology, Albert-Schweitzer-Strasse 33, D-48129 Muenster, Germany

(Received 3 November 1997; accepted for publication 30 June 1998)

Contrast-enhanced magnetic resonance imaging (MRI) of the breast is known to reveal breast cancer with higher sensitivity than mammography alone. The specificity is, however, compromised by the observation that several benign masses take up contrast agent in addition to malignant lesions. The aim of this study is to increase the objectivity of breast cancer diagnosis in contrast-enhanced MRI by developing automated methods for computer-aided diagnosis. Our database consists of 27 MR studies from 27 patients. In each study, at least four MR series of both breasts are obtained using FLASH three-dimensional (3D) acquisition at 90 s time intervals after injection of Gadopentetate dimeglumine (Gd-DTPA) contrast agent. Each series consists of 64 coronal slices with a typical thickness of 2 mm, and a pixel size of 1.25 mm. The study contains 13 benign and 15 malignant lesions from which features are automatically extracted in 3D. These features include margin descriptors and radial gradient analysis as a function of time and space. Stepwise multiple regression is employed to obtain an effective subset of combined features. A final estimate of likelihood of malignancy is determined by linear discriminant analysis, and the performance of classification by round-robin testing and receiver operating characteristics (ROC) analysis. To assess the efficacy of 3D analysis, the study is repeated in two-dimensions (2D) using a representative slice through the middle of the lesion. In 2D and in 3D, radial gradient analysis and analysis of margin sharpness were found to be an effective combination to distinguish between benign and malignant masses (resulting area under the ROC curve: 0.96). Feature analysis in 3D was found to result in higher performance of lesion characterization than 2D feature analysis for the majority of single and combined features. In conclusion, automated feature extraction and classification has the potential to complement the interpretation of radiologists in an objective, consistent, and accurate way. [S0094-2405(98)01509-0]

Key words: breast imaging, magnetic resonance imaging (MRI), computer-aided diagnosis, ROC analysis, contrast agent

I. INTRODUCTION

Breast cancer is a major cause of death among women in most western countries. Although mammography has demonstrated to be the most efficient tool for early detection of breast cancer, the technique may result in a missed fraction of cancers as high as 9%.[1] In addition, the fraction of lesions found by mammography that is sent to biopsy and proves to be malignant can be as low as 10%–20%.[2] Accurate examination of mammograms is particularly difficult in dense breasts, because lesions may be occluded by dense tissue. Consequently, complementary information by ultrasound or biopsy is often obtained.

Magnetic resonance imaging (MRI) is a promising complementary technique to mammography because of its inherent three-dimensional (3D) nature. In addition to possible improvement of diagnostic accuracy from dense breasts, MRI has shown superior potential for quantification of tumor volume, and detection of multifocal and multicentric disease.[3,4] These issues are of interest in the consideration of breast-conserving therapy. From current consensus, MR is particularly suited for specific problem cases, such as patients who have high risk of developing breast cancer, patients with implants, postoperative scars, or clinical evidence of breast cancer that cannot be detected by conventional diagnostic methods.[4–6]

MRI has become practically useful for breast imaging since the introduction of contrast agents that alter the spin-lattice (T_1) relaxation time.[7,8] Due to increased vascularity and capillary permeability of tumors,[9] contrast-enhanced MRI shows better distinction between lesions and normal tissue than conventional MRI alone. Nonetheless, contrast-enhanced MRI is known to enhance both malignant as well as some benign types of masses, thus compromising the specificity of the technique. In general, the sensitivity reported for diagnosis of breast cancer in MR images is larger than 90%,[10] but the reported specificity varies considerably and may be substantially lower.[6,10] The majority of these studies are solely based on enhancement as a function of

time. To improve the specificity without reducing the sensitivity, the morphology of enhancement has been studied as well.[11,12] Most of these techniques are, however, based on slice by slice assessment of the morphology in 2D. The performance is likely to improve when one takes full advantage of the 3D nature of the MR data.

An important aspect that may contribute to varying specificity is interobserver variation in the interpretation of the MR images. Nearly all studies presented to date are based on visual assessment by one or multiple radiologists. Mussurakis *et al.*[13] report significant variability in the assessment of lesions in MR data by human readers and stress the importance of standardized terminology. Heywang-Köbrunner *et al.*[6] indicate that differences in interpretation guidelines will influence the accuracy of contrast-enhanced MRI. Attempts to increase the objectivity of the interpretation have recently been reported using quantitative rating of features such as spiculation by a radiologist, followed by merging of these ratings using an interpretation model.[12] In this scheme, the classification stage is objective, but the rating of the features is still subjective to the interpretation of the radiologist.

Automated quantification and classification of features to discriminate between benign and malignant lesions has been pursued in other diagnostic areas such as mammography in the context of computer-aided diagnosis (CAD).[14] Several investigators have successfully developed methods for computerized detection[15–17] and computerized classification[18–20] for ultimate use in CAD as "second readers" for radiologists. In addition to objective analysis, computerized analysis can take full advantage of information across slices in 3D data sets which is difficult to assess visually from individual images.

The aim of this study is to increase the objectivity of breast cancer diagnosis in contrast-enhanced MRI. This aim is pursued by automated extraction of features that quantify spatial properties of contrast enhancement in 3D, and by merging different features into an estimate of malignancy using automated classification. The ultimate objective of this feasibility study is to reduce the number of biopsies of benign lesions and to increase the sensitivity for cancer cases.

II. MATERIAL AND METHODS

A. Image and patient data

The images in this study were obtained using fast low-angle shot (FLASH) 3D acquisition at field strength of 1.0 T (Siemens Impact, Siemens, Erlangen, Germany). The acquisition parameters were: $TR = 14.0$ ms, $TE = 7.0$ ms, and flip angle of 25°. Fat suppression was not employed. The patients were scanned in prone position using a standard double-breast coil. In total, 141 preoperative MR series were acquired from 27 patients. Each series contains 64 coronal slices with a typical field of view of 32×16 cm^2. Each slice contains 256×256 pixels of 1.25×1.25 mm^2 and has a typical thickness of 2 mm. There are no gaps in between the slices. Gadopentetate dimeglumine (Gd-DTPA) contrast agent was injected intravenously by power injection after acquisition of the precontrast MR series. At least four series were taken per patient at 90 s intervals. Figure 1 shows an example of a dynamic MR sequence on a single slice through a malignant lesion.

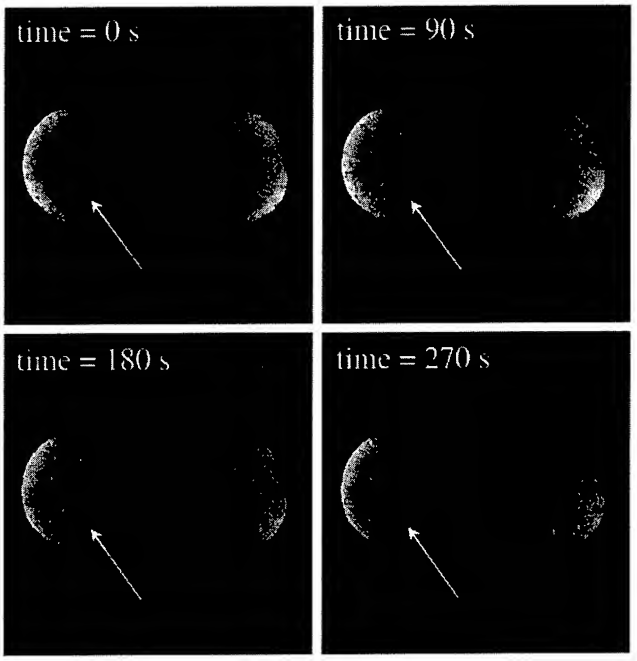


FIG. 1. Example of contrast enhancement in a malignant lesion, illustrated on a dynamic series of a single MR slice. Note the irregular "donut" shape of the lesion (arrow) as it enhances in time before merging with the background.

The database in this study contains 28 lesions: 13 benign and 15 malignant masses. Histology in 27 out of 28 lesions was confirmed by open excisional biopsy, one case was benign based on core biopsy and four-year follow up. The distribution of the size of the lesions is shown in Fig. 2. The relative position of the lesions in the breast varies, some are close to the skin and near the chest wall. Benign masses include fibroadenoma (6/13), papilloma (2/13), and benign mastopathy (5/13). Malignant cases include papillary (1/15),

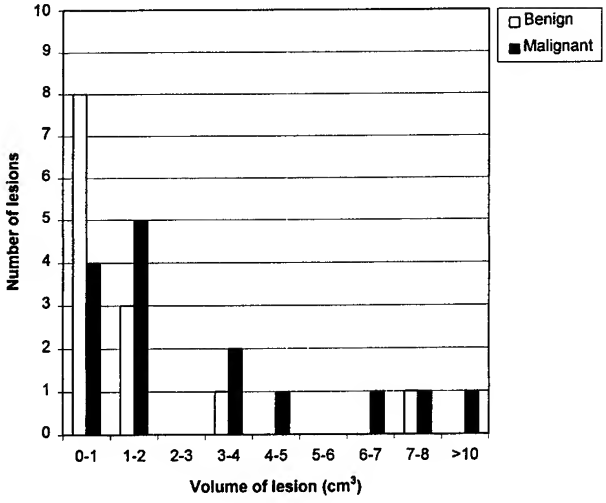


FIG. 2. Distribution of the size of benign and malignant lesions in our database.

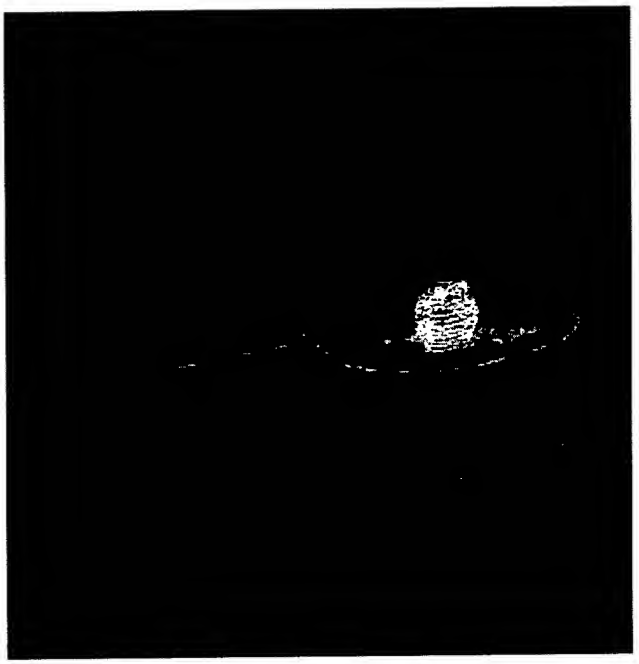


FIG. 3. Visualization of lesion shape, size, and orientation in 3D with respect to reference landmarks. The MR data is acquired in slices in coronal orientation, but the computerized analysis of the lesion is performed in all three dimensions simultaneously.

tubular (2/15), and medullary carcinoma (1/15), invasive lobular (3/15) and ductal (7/15) cancer, and ductal carcinoma *in situ* (DCIS) (1/15).

B. Feature extraction

In this preliminary study, suspect masses were delineated by a radiologist (U.B.) experienced in MR-mammography and blinded concerning the histological diagnosis. This segmentation was performed in the subtraction images (postcontrast–precontrast) by contouring the enhanced tumor area in each slice that intersected the lesion. All available subtraction images were used for this purpose. As an additional reference, the radiologist had access to the original (nonsubtracted) MR images as well. All other stages of the scheme, though, as described below, are fully automated.

The proposed strategy for computerized analysis of dynamic MR data in 3D consists of two consecutive stages: Feature extraction and classification. The feature extraction stage is aimed at quantification of spatial properties of enhancement in suspicious lesions. Feature extraction has two parts: Extraction of the breast volume, and quantification of spatial properties. Although the MR data is obtained in slices, the analysis of the lesion is performed in 3D, taking all directions into account (Fig. 3). The volume of the breast is extracted from the MR data by global segmentation of pixel values at a threshold that maximizes the interclass variance between two pixel-value regions.[21] All slices in the data contribute to the computation of a single threshold value. The result of the segmentation is a 3D binary mask in which breast voxels are labeled with value "1," and background voxels with value "0." Remaining gaps in the mask are removed by morphological closing operations:[22] Morphological erosion[22] is employed to remove an empirically established margin of two voxels from the external surface of the breast mask. This step is required to avoid strong voxel-value gradients near the borders of the breast to be included in the computation of gradient-based lesion features. Subsequent computation of features in the original MR data is restricted to voxels that have value "1" at the corresponding locations in the breast mask.

Features investigated in this study concern the inhomogeneity of uptake in the lesion [Eqs. (1) and (2)], sharpness of the lesion margin [Eqs. (3) and (4)], analysis of the shape of the lesion [Eqs. (5) and (6)], and radial gradient analysis [Eq. (7)].

If the set of voxel values in the lesion at time frame "i" is given by $F_l(\mathbf{r},i)$, where vectors $\mathbf{r}$ point to the lesion, and index "i" runs from frame 0 (i.e., the frame before injection of contrast) to $M-1$ (where M is the total number of time frames), then the inhomogeneity of contrast uptake is characterized by two features which are defined by

$$\max_{i=0,\ldots,M-1}\left\{\frac{\text{variance}_{\mathbf{r}}F_l(\mathbf{r},i)}{\text{variance}_{\mathbf{r}}F_l(\mathbf{r},0)}\right\}, \tag{1}$$

referred to here as variance of uptake

and

$$\min_{i=0,\ldots,M-2}\left\{\frac{\text{variance}_{\mathbf{r}}F_l(\mathbf{r},i)}{\text{variance}_{\mathbf{r}}F_l(\mathbf{r},i+1)}\right\}, \tag{2}$$

referred to here as change in variance of uptake,

where $\text{variance}_{\mathbf{r}}F_l(\mathbf{r},i)$ denotes the computation of the variance of the voxel values at all $\mathbf{r}$ in the lesion at fixed time frame "i."

The sharpness of the lesion margins is characterized by two features as well. The first feature is given by

$$\max_{i=0,\ldots,M-1}\left\{\frac{\text{mean}_{\mathbf{r}}\|\nabla[F_m(\mathbf{r},i)-F_m(\mathbf{r},0)]\|}{\text{mean}_{\mathbf{r}}F_m(\mathbf{r},i)}\right\}, \tag{3}$$

referred to here as margin gradient,

where $\nabla[F_m(\mathbf{r},i)-F_m(\mathbf{r},0)]$ denotes the set of voxel-value gradients at the margin of the suspect lesion in the difference images of time frame "i" and precontrast frame "0." Thus, the sharpness of the uptake of contrast is computed at the lesion margin. The range of vectors $\mathbf{r}$ in F_m is limited to a shell—three voxels thick—centered on the surface of the lesion. The shell is employed to account for small inaccuracies that may occur in the delineation of the lesion outlines.

The second feature related to margin sharpness is defined by

$$\frac{\text{variance}_{\mathbf{r}}\|\nabla[F_m(\mathbf{r},i)-F_m(\mathbf{r},0)]\|}{[\text{mean}_{\mathbf{r}}F_m(\mathbf{r},i)]^2}, \tag{4}$$

referred to here as variance of margin gradient,

and is only computed from the subtraction frames of "i" and "0" where the margin gradient [Eq. (3)] is maximum.

In Eqs. (3) and (4), computation of the spatial gradient is accomplished in 3D by convolution with the components of a $3\times3\times3$ Sobel filter[23] in three orthogonal directions. Note that this approach takes information on lesion margins across slices into account.

Circularity of the shape of the lesion in 3D is given by

$$\frac{\text{volume of lesion within sphere of effective diameter}}{\text{volume of lesion}}, \tag{5}$$

and irregularity in 3D by

$$1-\frac{\pi\cdot\text{effective diameter}^2}{\text{surface of lesion}}, \tag{6}$$

where the effective diameter is defined by

$$2\cdot\sqrt[3]{\frac{3\cdot\text{volume of lesion}}{4\pi}}.$$

The volume and the surface of the lesion are estimated from the contours of the segmented masses. For this purpose, a set of binary images is created in which the pixels at and enclosed by the contours are set to value "1" (object pixels), and remaining pixels to value "0" (background pixels). The volume of the lesion is determined by multiplying the number of object pixels with the volume of one voxel in world coordinates. The surface of the lesion is computed by combining the set of 2D binary images into a 3D binary representation of the lesion. Next, the faces of the object voxels that are exposed to background voxels in the 3D binary volume are identified by examining the value of the neighboring voxels in the x, y, and z directions. The face of an object voxel is exposed to the background if the neighboring voxel has value "0." The surface of the lesion is subsequently determined by calculating the sum of the areas of the faces exposed to the background in world coordinates. Note that the circularity and irregularity in 3D are computed from the volume and the surface of the lesion in world coordinates—rather than in voxel coordinates—to account for the differences between pixel size and slice thickness (i.e., the anisotropic voxel shapes).

Radial gradient analysis is based on examination of the angles between voxel-value gradients and lines intersecting a single point near the center of the suspect lesion (i.e., lines in radial directions). Radial gradient values are given by the dot product of the gradient direction and the radial direction. The histogram of radial gradient values—quantifying the frequency of occurrence of the dot products in a given region of interest (Fig. 4)—is called the radial gradient histogram (RGH). Analysis of the RGH yields

$$\max_{i=0,\ldots,M-1}\{\text{variance}_{p>0}\, H(p)\}, \tag{7}$$

referred to here as the variance of RGH values.

In this relationship, $H(p)$ denotes the normalized RGH, and variable p is given by the normalized dot product

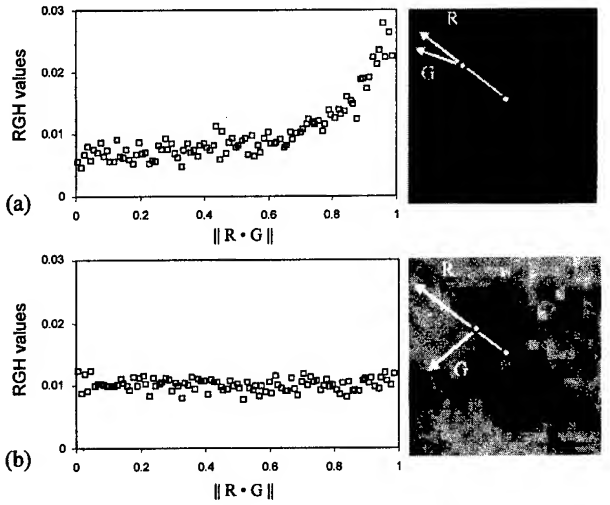


FIG. 4. The radial gradient histogram (RGH) of a volume of interest (VOI) with a benign lesion (a) and a malignant lesion (b). Shown are images of representative cross sections through the lesions. The radial vector (R) originates in the center of the VOI. The gradient vector (G) indicates the local direction of the voxel-value gradient. The RGH maps the dot product of R and G against the frequency of occurrence (RGH values). In benign lesions, R and G tend to point in comparable directions within the VOI, yielding a peak in the RGH around 1.0 (a). Malignant lesions typically extend in less spherical patterns resulting in a flat RGH (b). The variance of RGH values is used to quantify the flatness of the RGH.

$$p=\frac{|\nabla[F_b(\mathbf{r},i)-F_b(\mathbf{r},0)]\cdot(\mathbf{r}-\mathbf{r}_c)|}{\|\nabla[F_b(\mathbf{r},i)-F_b(\mathbf{r},0)]\|\cdot\|\mathbf{r}-\mathbf{r}_c\|},$$

where $\nabla[F_b(\mathbf{r},i)-F_b(\mathbf{r},0)]$ indicates the set of voxel-value gradients in a rectangular box of interest at the subtracted time frames "i" and "0." The box encompasses the suspect lesion with an additional margin of three voxels along all sides. Vector $\mathbf{r}_c$ points to the center of this rectangular box. Thus, $\mathbf{r}_c$ generally will not (and does not need to) point to the exact center of the lesion. This aspect will be reviewed in more detail in the Discussion section.

In essence, above equations quantify the observation that malignant lesions take up contrast agent in a less homogeneous pattern than benign masses, have less sharp boundaries and are more irregularly shaped. "Circularity" quantifies how well the lesion conforms to a spherical shape [Eq. (5)], and "irregularity" indicates the roughness of the surface of the lesion [Eq. (6)]. Radial gradient analysis was previously applied to mammograms to quantify spiculation of projected masses.[18] The analysis provides a measure that indicates how well the image structures in a region of interest (ROI) extend in a radial pattern originating from the center of the ROI. Round and well-defined masses produce different measures than irregular and spiculated lesions. In the current study, radial gradient analysis is extended to 3D. The feature "Variance of RGH values" quantifies how well the image structures in a volume of interest (VOI) extend in a spherical pattern originating from the center of the VOI (Fig. 4).

With the exception of circularity and irregularity, which are computed from the coordinates of the segmented lesions, all other features are extracted from the data at each available

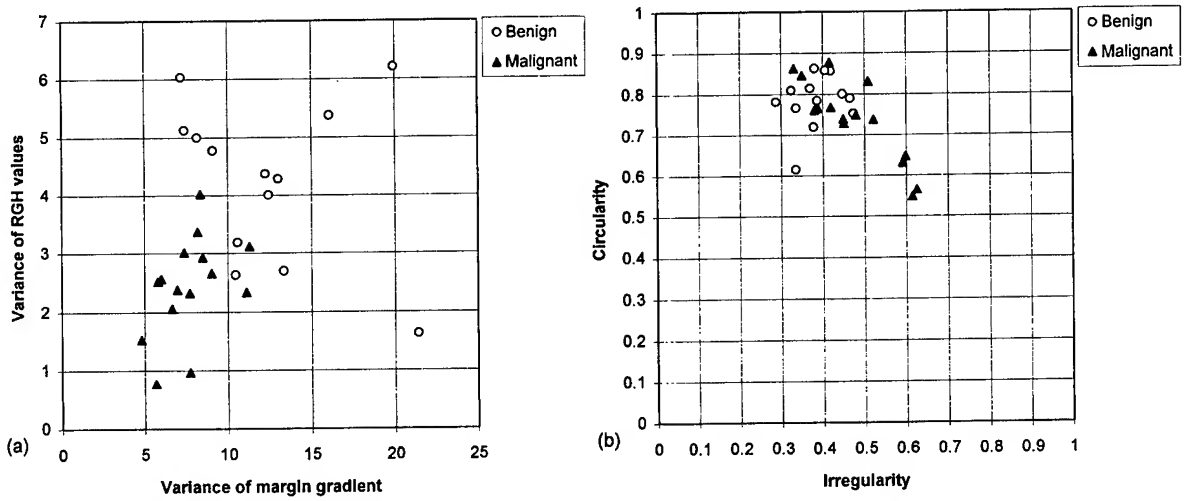


FIG. 5. Relationships of various features for the database of benign and malignant lesions (a) margin gradient analysis and radial gradient analysis, (b) shape analysis.

time frame and combined by minimum or maximum operations, such as described by Eqs. (1)–(4) and (7).

To assess the efficacy of 3D analysis in comparison with conventional 2D techniques, the feature extraction and analysis procedures were repeated, although in 2D, on a single representative slice through the middle of the lesions.

C. Feature selection and classification

From the total set of seven features, stepwise multiple regression[24] produced a selection that performs efficiently in distinguishing between benign and malignant lesions. The technique involves adding and removing features to obtain a limited subset that provides statistically significant separation in the estimated likelihood of malignancy. Linear discriminant analysis[25] is employed to estimate this likelihood of malignancy from single or combined features.

D. Evaluation

The performance of the computerized method in classification (distinguishing between benign and malignant lesions) is quantified by receiver operating characteristics (ROC) analysis.[26] In particular the area under the ROC curve (A_z)—which maps the fraction of false positives to the fraction of true positives—is used as a measure of performance in this study. Sensitivity is defined as the true-positive fraction, specificity as one minus the false–positive fraction. The area under the ROC curve at true-positive fractions larger than 0.9 (partial A_z) is employed to rate the performance of computerized analysis at high sensitivity levels.[27]

The general performance of the computerized method is estimated by round-robin testing[28] on our current database. This "leave-one-out" technique involves estimating the likelihood of malignancy from all cases but one, testing classification on that single case, and repeating the procedure until each case has been tested individually.

III. RESULTS

All features investigated in this study show potential for distinguishing between benign and malignant lesions (Fig. 5, Table I). As expected, benign masses were found to extend more along spherical patterns than malignant lesions, and the margins of benign masses were found to be sharper on average than the margins of malignant lesions. An interesting observation is, however, that the variance of sharpness along the margin of the lesions is larger on average for benign than for malignant masses [Fig. 5(a)]. A possible explanation is offered in the discussion section. Less surprising was the result that some malignant lesions tend to be more irregularly shaped than benign masses [Fig. 5(b)]. Circularity was, however, not found to be a strong feature to distinguish between

TABLE I. Area under the ROC curves (A_z) using 2D and 3D analysis of individual and combined features. The standard deviations (1 SD) are shown in parentheses.

Feature	A_z (2D)	A_z (3D)
Inhomogeneity of uptake		
Variance of uptake	0.54 (0.11)	0.72 (0.11)
Change in variance of uptake	0.59 (0.11)	0.77 (0.10)
Sharpness		
Margin gradient	0.83 (0.07)	0.88 (0.07)
Variance of margin gradient	0.71 (0.10)	0.86 (0.07)
Shape		
Circularity	0.67 (0.10)	0.65 (0.10)
Irregularity	0.66 (0.10)	0.80 (0.08)
Radial gradient analysis		
Variance of RGH values	0.80 (0.08)	0.88 (0.07)
Combinations of features		
Variance of RGH values and margin gradient	0.87 (0.11)	0.92 (0.05)
variance of RGH values and variance of margin gradient	0.86 (0.08)	0.96 (0.03)

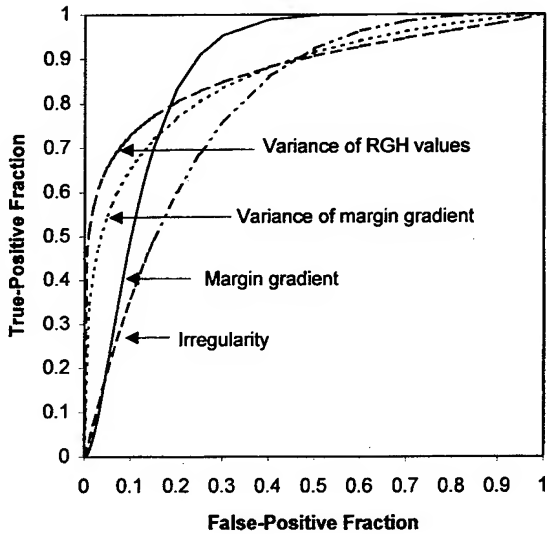


FIG. 6. ROC curves showing the performance of the best single features in the task of distinguishing between benign and malignant lesions. (RGH=Radial gradient histogram.)

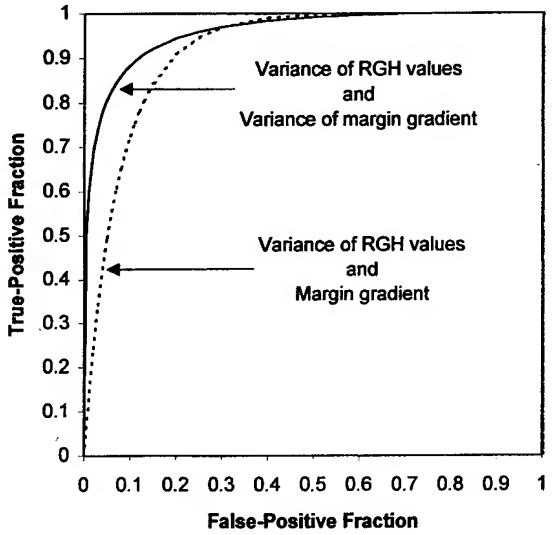


FIG. 7. ROC curves showing the performance of effective combined features in the task of distinguishing between benign and malignant lesions (using round-robin testing). (RGH=Radial gradient histogram.)

benign and malignant in our database (Table I). Using single features only, the highest performances were obtained with radial gradient analysis (A_z=0.88), sharpness (A_z=0.88 and 0.86), and shape analysis of the lesion (A_z=0.80). The corresponding ROC curves are shown in Fig. 6. For operation at sensitivities larger than 0.9, the highest performance was achieved with the "margin gradient" feature, yielding a partial A_z value of 0.70. Note that although the "margin gradient" and "variance of margin gradient" have comparable A_z values (Table I), their ROC curves are differently shaped (Fig. 6). The curve of the "margin gradient" feature is steeper, indicating that higher specificity can be achieved at high sensitivity. In addition, "variance of RGH values" and "margin gradient" have comparable A_z values, although the shapes of their ROC curves differ. The "margin gradient" feature was found to perform better at higher sensitivity than the "variance of RGH values" (Fig. 6).

Stepwise multiple regression at a confidence region of 95% resulted in combination of two features: "Variance of RGH values" [eq. (7)] and "variance of margin gradient" [eq. (4)]. Their combined performance in distinguishing between benign and malignant lesions resulted in an A_z value of 0.96. The corresponding ROC curve is shown in Fig. 7. Some combinations that have more than two features yielded slightly better results, but this increase in performance was not found to be statistically significant.

Based on round-robin testing on our database, the computerized scheme achieved 77% specificity (10/13) at 100% sensitivity. Note that all lesions in our database had been biopsied. In other words, all benign lesions were basically misclassified before biopsy, whereas the computerized method misclassified only 3 of the 13 benign lesions without misclassifying any malignant cases. These preliminary results indicate that the computerized method has the potential to reduce the number of biopsies of benign lesions.

Although statistical significance of the differences in performance between 3D and 2D analysis of features could not be ascertained given the current size of our database, a consistent superior performance using 3D analysis was found for nearly all single and all combined features (Table I).

IV. DISCUSSION

Automated extraction of mathematically defined features in 3D yields encouraging results in distinguishing benign from malignant lesions (A_z=0.96). A_z values of classification achieved by manual rating of features by radiologists have been reported[12] to be around 0.86. Direct comparison with results reported in the literature is, however, difficult due to differences in database used. In addition, most reports indicate only a single operating point for sensitivity and specificity on the ROC curve. In our study, the operating point must be tuned to a desired trade-off between sensitivity and specificity. This trade-off would be determined clinically by a cost-benefit analysis. In our current study, a sensitivity of 100% can be obtained at specificity of 77%. At this operating point, all malignant lesions, including one case of DCIS, are successfully identified as malignant. One of the next steps of research is to study the accuracy of breast cancer diagnosis done by radiologists when assisted by the automated technique as a second opinion.

We found that the margins of benign lesions are typically sharper in appearance than the margins of malignant lesions, which is consistent with observations from other studies. When the average sharpness of the lesion margins is high, deviations from this average caused by anatomical morphology, partial volume effect, or inaccuracies in segmentation of the lesion, will result in higher variance values than would occur with a small average sharpness of lesion margins. Thus, the variation of sharpness along the margin is expected

to be larger for benign than for malignant cases. Our preliminary results indicate that this observation yields good potential to discriminate between benign and malignant lesions [Fig. 5(a), Table I].

In the current study, we found that spatial features are effective to distinguish between benign and malignant lesions, in particular the combination of radial gradient analysis and analysis of margin sharpness. Most studies of contrast-enhanced MRI of the breast are based on analysis of temporal features of uptake only—such as speed of uptake—and report varying specificity. It is possible, however, that consistent high performance can be obtained from temporal features when the temporal resolution of the data is high. Boetes *et al.*[29] report encouraging results from temporal features in data obtained at high temporal resolution at the expense of spatial resolution. Physiological models of contrast uptake and washout have also been applied to increase the specificity of the diagnosis from temporal features, e.g., Tofts *et al.*[30] The preliminary results from our study indicate that good distinction between benign and malignant lesions can be obtained from spatial features without extremities in temporal or spatial resolution. An important aspect is, however, 3D acquisition and analysis of the spatial features. The results from the current study indicate that it is beneficial to analyze spatial features in 3D rather than in 2D to distinguish between benign and malignant lesions.

In addition to differences in image acquisition, other aspects may also influence the specificity of the diagnosis. Differences in bolus size and particular hemodynamic characteristics of each patient as well as hormonal factors, may cause variable enhancement.[4,6] Image artifacts can be caused by inhomogeneity of the magnetic field and by patient motion. To reduce the effect of some of these aspects, features have been normalized within or across time frames in the same examination. Gradient artifacts caused by inhomogeneity of the magnetic field typically occur at much lower spatial frequency than the lesion margins, and were found to be of less importance in this study. Patient movement is estimated to be about 2 mm on average in our data set. Because the voxel dimensions are $1.25\times1.25\times2.0\ \text{mm}^3$, the motion causes some blurring of the lesions rather than an actual displacement of image structures. To avoid image artifacts due to motion of the heart, the MR slices were obtained in coronal orientation. Different slice thickness may also result in differences in performance. To take the anisotropic voxel shapes in the MR data into account, features related to the shape of the lesion are calculated in world coordinates, rather than in voxel coordinates. Lesion sharpness is calculated, however, using a $3\times3\times3$ Sobel filter, which does not account for the anisotropic voxel shape. An intuitive approach would be to resample the MR voxels into a uniform coordinate grid by linear interpolation. Such linear modeling of the discontinuity at the edges of the lesions may, however, lead to underestimation of sharp lesion margins, thus compromising the benefit of the correction. In addition, the approach does not take other effects of deviating slice thickness into account, such as differences in partial volume effect. These aspects are topic of future study.

Automated segmentation of the lesions is likely to further improve the objectivity of diagnosis. Techniques for this purpose have recently been investigated, e.g., Lucas-Quesada *et al.*,[31] and will remain a subject of future research. In our study, the sharpness-related features are computed in a shell around the indicated outline of the tumor to account for small inaccuracies in the segmentation. Radial gradient analysis does not require accurate delineation of the margins of the lesion: The region of interest is a rectangular box positioned roughly around the lesion in the subtracted images. Spurious gradients associated with background noise are expected to be randomly distributed with respect to the radial directions of the noise voxels, thus adding a small constant offset to the RGH values. Consequently, a region of interest somewhat larger than the size of the actual lesion is not expected to have much effect on the variance of the RGH values. The radial lines intersect the center of the rectangular bounding box positioned around the lesion. This intersection point will usually not coincide with the exact center of the lesion. Nevertheless, since the shape of the lesions is generally not perfectly symmetrical and regular, the center of the lesion is expected to vary with the shape of the lesions in a similar way as the center of the rectangular bounding box is expected to vary with the shape of the lesions. Consequently, neither definition of the center is expected to yield superior performance compared to the other.

It is likely that the ability to distinguish between benign and malignant lesions will decrease for smaller tumor sizes. The smallest lesion in our database has a volume of 0.1 cm^3 (benign case) and was correctly classified at no loss of malignant cases. At this operating point, 3 of the 13 benign lesions were incorrectly classified as malignant. These lesions had volumes of 0.2, 0.9, and 3.4 cm^3, respectively (papilloma and benign mastopathy). Other benign lesions with similar histology and volumes were, however, correctly classified, as well as all malignant lesions, which have sizes ranging from 0.1 cm^3 to larger than 10 cm^3 (Fig. 2). In conclusion, our database did not show an obvious correlation between accuracy of the performance of the computerized diagnosis and lesion size, nor between accuracy and histology. Evaluation of the computerized analysis technique on larger databases is, however, required, and may warrant the use of more advanced classification methods, such as artificial neural networks.

Once a mass is suspected to be malignant, localization for accurate biopsy is a next step. Techniques for MRI-directed biopsy are being developed and evaluated.[32,33] The 3D nature of the MR data may allow useful complementary information to visualize the size and shape of the lesion as well as its location relative to the nipple position and pectoralis muscle (Fig. 3).

V. CONCLUSIONS

A technique aimed at computer-aided diagnosis of suspect lesions in contrast-enhanced MRI of the breast has been developed to increase the objectivity of breast cancer diagnosis. Initial results of analysis of spatial features in 3D indicate good accuracy of classification ($A_z=0.96$), and higher per-

formance than analysis in 2D for the majority of single and combined features. Consequently, automated extraction of features that quantify the spatial properties of contrast enhancement has the potential to complement the interpretation of radiologists in an objective, consistent and accurate way. In addition, the computerized analysis technique shows potential to reduce the fraction of biopsies of benign lesions.

ACKNOWLEDGMENTS

The authors thank K. Doi, G. S. Karczmar, Z. Huo, H. Al-Hallaq, C. E. Metz, and M. Kupinski for helpful discussions. This work was supported in part by the Netherlands Cancer Institute (NKI/AvL), U.S. Army Grants No. DAMD17-96-1-6058 and No. DAMD17-98-1-8194, and USPHS Grants No. CA 24806 and No. RR 11459. M. L. Giger is a shareholder of R2 Technology, Inc., Los Altos, CA. It is the policy of the University of Chicago that investigators disclose publicly actual or potential significant financial interests that may appear to be affected by the research activities.

[a]Author to whom correspondence should be addressed. Also at: The Netherlands Cancer Institute/Antoni van Leeuwenhoek Huis, Radiotherapy Department, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands; Electronic mail: kengil@nki.nl

[b]Currently at: University Clinics Charite, Department of Radiology, Schumannstrasse 20121 10117 Berlin, Germany.

[1]L. H. Baker, "Breast cancer detection demonstration project: five year summary report," Cancer **32**, 194–225 (1982).

[2]L. W. Bassett and R. H. Gold, *Breast Cancer Detection: Mammography and Other Methods in Breast imaging* (Grune & Stratton, New York, 1987).

[3]P. L. Davis, M. J. Staiger, K. B. Harris, M. A. Gannot, J. Klementaviciene, K. S. McCarty, Jr., and H. Tobon, "Breast cancer measurements with magnetic resonance imaging, ultrasonography, and mammography," Breast Cancer Res. Treatment **37**, 1–9 (1996).

[4]S. E. Harms, "MRI in breast cancer diagnosis and treatment," Curr. Probl. Diagn. Radiol. **25**, 193–215 (1996).

[5]S. E. Harms, D. P. Flamig, K. L. Hesley, M. D. Meiches, R. A. Jensen, W. P. Evans, D. A. Savino, and R. V. Wells, "MR imaging of the breast with rotating delivery of excitation off resonance: clinical experience with pathologic correlation," Radiology **187**, 493–501 (1993).

[6]S. H. Heywang-Köbrunner, P. Viehweg, A. Heinig, and Ch. Küchler, "Contrast-enhanced MRI of the breast: accuracy, value, controversies, solutions," Eur. J. Radiol. **24**, 94–108 (1997).

[7]S. H. Heywang, D. Hahn, H. Schmid, I. Krischke, W. Eiermann, R. Bassermann, and J. Lissner, "MR imaging of the breast using gadolinium-DTPA," J. Comput. Assist. Tomogr. **10**, 199–204 (1986).

[8]W. A. Kaiser and E. Zeitler, "MR imaging of the breast: fast imaging sequences with and without Gd-DTPA," Radiology **170**, 681–686 (1989).

[9]D. Revel, R. Brasch, H. Paajanen, W. Rosenau, W. Grodd, B. Engelstad, P. Fox, and J. Winkelhake, "Gd-DTPA contrast enhancement and tissue differentiation in MR imaging of experimental breast carcinoma," Radiology **158**, 319–323 (1986).

[10]F. Kelcz and G. Santyr, "Gadolinium-Enhanced breast MRI," Crit. Rev. Diagn. Imaging **36**, 287–338 (1995).

[11]A. H. Adams, J. R. Brookeman, and M. B. Merickel, "Breast lesion discrimination using statistical analysis and shape measures on magnetic resonance imagery," Comp. Med. Imaging Graphics **5**, 339–349 (1991).

[12]L. W. Nunes, M. D. Schnall, S. G. Orel, M. G. Hochman, C. P. Langlotz, C. A. Reynolds, and M. H. Torosian, "Breast MR imaging: Interpretation Model," Radiology **202**, 833–841 (1997).

[13]S. Mussurakis, D. L. Buckley, A. M. Coady, L. W. Turnbull, and A. Horseman, "Observer variability in the interpretation of contrast enhanced MRI of the breast," Br. J. Radiol. **69**, 1009–1016 (1996).

[14]M. Giger and H. MacMahon, "Image processing and computer-aided diagnosis," Radiol. Clin. N. Am. **34**, 565–596 (1996).

[15]H. P. Chan, K. Doi, S. Galhotra, C. J. Vyborny, H. MacMahon, and P. M. Jokich, "Image feature analysis and computer-aided diagnosis in digital radiography: 1. Automated detection of microcalcifications in mammography," Med. Phys. **14**, 538–548 (1987).

[16]W. P. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for spiculated lesions," Radiology **191**, 331–337 (1994).

[17]N. Karssemeijer, "Recognition of stellate lesions in digital mammograms," *Digital Mammography* (Elsevier Science B.V., Amsterdam, The Netherlands, 1994).

[18]Z. Huo, M. L. Giger, C. J. Vyborny, U. Bick, P. Lu, D. E. Wolverton, and R. A. Schmidt, "Analysis of spiculation in the computerized classification of mammographic masses," Med. Phys. **22**, 1569–1579 (1995).

[19]Y. Jiang, R. M. Nishikawa, D. E. Wolverton, C. E. Metz, M. L. Giger, R. A. Schmidt, C. J. Vybourny, and K. Doi, "Automated feature analysis and classification of malignant and benign clustered microcalcifications," Radiology **198**, 671–678 (1996).

[20]R. M. Rangayyan, N. El-Faramawy, J. E. L. Desautels, and O. A. Alim, "Discrimination between benign and malignant breast tumors using a region-based measure of edge profile acutance," *Digital Mammography* (Elsevier Science B.V., Amsterdam, The Netherlands, 1996).

[21]N. Otsu, "A threshold selection method from gray-value histograms," IEEE Trans. Syst. Man Cybern. **9**, 62–66 (1979).

[22]J. Serra, *Image Analysis and Mathematical Morphology*, 2nd ed. (Academic, San Diego, California, 1988), Vol. 1.

[23]R. C. Gonzales and P. Wintz, *Digital Image Processing*, 2nd ed. (Addison-Wesley, Reading, Massachusetts, 1987).

[24]J. Neter, W. Wasserman, and M. H. Kutner, *Applied Linear Statistical Models Regression, Analysis of Variance, and Experimental Designs*, 2nd ed. (Irwin, Homewood, 1985).

[25]R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 3rd ed. (Prentice-Hall, Englewood Cliffs, New Jersey, 1992).

[26]C. E. Metz, "Some practical issues of experimental design and data analysis in radiological ROC studies," Invest. Radiol. **24**, 234–245 (1989).

[27]Y. Jiang, C. E. Metz, and R. M. Nishikawa, "A receiver operating characteristic partial area index for highly sensitive diagnostic tests," Radiology **201**, 745–750 (1996).

[28]G. Gong, "Cross-validation, the Jackknife, and the Bootstrap: Excess Error Estimation in Forward Logistic Regression," J. Am. Stat. Assoc. **81**, 108–113 (1986).

[29]C. Boetes, J. O. Barentsz, R. D. Mus, R. F. van der Sluis, L. J. T. O. van Erning, J. H. C. L. Hendriks, R. Holland, and S. H. J. Ruys, "MR characterization of suspicious breast lesions with a Gadolinium-enhanced TurboFLASH subtraction technique," Radiology **193**, 777–781 (1994).

[30]P. S. Tofts, B. Berkowitz, and M. D. Schnall, "Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model," Magn. Reson. Med. **33**, 564–568 (1995).

[31]F. A. Lucas-Quesada, U. Sinha, and S. Sinha, "Segmentation strategies for breast tumors from dynamic MR images," J. Magn. Reson. Imaging **6**, 753–763 (1996).

[32]S. H. Heywang-Köbrunner, A. T. Huynh, P. Viehweg, W. Hanke, H. Requardt, and I. Paprosch, "Prototype breast coil for MR-guided needle localization," J. Comput. Assist. Tomogr. **18**, 876–881 (1994).

[33]S. G. Orel, M. D. Schnall, R. W. Newman, C. M. Powell, M. H. Torosian, and E. F. Rosato, "MR imaging—guided localization and biopsy of breast lesions: initial experience," Radiology **193**, 97–102 (1994).

Submitted to Academic Radiology
Confidential until publication

Computerized analysis of lesions in ultrasound images of the breast

Maryellen L. Giger, Ph.D.

Hania Al-Hallaq, B.S.

Zhimin Huo, Ph.D.

Catherine Moran, B.A.

Dulcy E. Wolverton, M.D.

Chun Wai Chan, M.S.

Weiming Zhong, M.S.

Corresponding author
Maryellen L. Giger, Ph.D.
Kurt Rossmann Laboratories for Radiologic Image Research
Department of Radiology, MC 2026
5841 S. Maryland Ave.
Chicago, IL 60637
phone: 773-702-6778
fax: 773-702-0371
e-mail: m-giger@uchicago.edu

Running title: Computer analysis of breast ultrasound

Abstract

Rationale and Objectives. Breast sonography is not routinely used to distinguish benign from malignant solid masses because of the considerable overlap in their sonographic appearances. We are developing computerized analyses of breast lesions in ultrasound images to aid in the discrimination between malignant and benign lesions.

Materials and Methods. We extracted and calculated features related to lesion margin, shape, homogeneity (texture) and the nature of the posterior acoustic attenuation pattern in ultrasound images of the breast. Our database contained 184 digitized ultrasound images from 58 patients with 78 lesions. Benign lesions were confirmed by biopsy, cyst aspiration, or image interpretation alone, while malignant lesions were confirmed by biopsy. ROC analysis was used to study the performance of the various individual features and the output from linear discriminant analysis in distinguishing benign from malignant lesions.

Results. From ROC analysis, the feature characterizing the margin yielded A_Z values of 0.85 and 0.75, in the task of distinguishing between benign and malignant lesions in the entire database and in an equivocal database, respectively. The "equivocal" database contained lesions that had been proven to be benign or malignant by either cyst aspiration or biopsy. Linear discriminant analysis round-robin runs yielded A_Z values of 0.94 and 0.87 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively.

Conclusion. Results indicate that computerized analysis of ultrasound images has the potential to increase the specificity of breast sonography.

Keywords: computer-aided diagnosis, ultrasound imaging, breast imaging, differential diagnosis, computer vision, artificial intelligence

Introduction

Breast cancer is a leading cause of death in women, causing an estimated 44,000 deaths per year (1). Mammography is the most effective method for the early detection of breast cancer, and it has been shown that periodic screening of asymptomatic women does reduce mortality (2-4). Many breast cancers are detected and referred for biopsy on the basis of a radiographically detected mass lesion or cluster of microcalcifications. Although general rules for the differentiation between benign and malignant mammographically identified breast lesions exist (5, 6), considerable misclassification of lesions occurs with the current methods. On average, less than 30% of masses referred for surgical breast biopsy are actually malignant (7).

Breast sonography is used as an important adjunct to diagnostic mammography and is typically performed to evaluate palpable and mammographically identified masses in order to determine their cystic vs. solid natures. The accuracy of ultrasound has been reported to be 96-100% in the diagnosis of simple benign cysts (8). Masses so characterized do not require further evaluation. Ultrasound has not been used for screening purposes due to relatively high false-negative and false-positive rates. However, ultrasound is being evaluated as a potential method to screen women with dense breasts (9). While some centers are successful at visual diagnosis of benign from malignant masses using ultrasound, most facilities are unable to rely on breast ultrasound to avoid biopsy because of the considerable overlap in their sonographic appearances. With the advent of modern high-frequency transducers that have improved spatial and contrast resolution, a number of sonographic features have emerged as potential indicators of malignancy, while other features are typical for benign masses (10,11). Benign features include hyperechogenicity, ellipsoid shape, mild lobulation, and a thin, echogenic pseudocapsule. Malignant features include spiculation, angular margins, marked hypoechogenicity, posterior acoustic shadowing, and depth-to-width ratio greater than 0.8.

Stavros, et al. used various features to characterize masses as benign, indeterminate, and malignant (12). Their classification scheme had a sensitivity of 98.4% and a negative predictive value of 99.5%. However, the sonographic evaluation described by these investigators is much more extensive and complex than is usually performed at most breast imaging centers. Ultrasound examination is a notoriously operator-dependent modality, and until these encouraging results are corroborated through additional studies by other investigators, it is unclear how widely applicable or reliable such sonographic classification schemes will be.

Computer-aided techniques have been applied to the color Doppler evaluation of breast masses with promising results (13). However, color Doppler imaging is a technique which focuses only upon the vascularity of lesions. Since not all sonographically visible cancers have demonstrable neovascularity and benign lesions can be vascular, the sensitivity and specificity of this technique is inherently somewhat limited. These limitations have been demonstrated in power Doppler imaging of solid breast masses (14).

Comprehensive summaries of investigations in the field of mammography computer-aided diagnosis (CAD) have been published (15,16). In the 1960s and 70s, several investigators attempted to analyze mammographic abnormalities with computers (17-24). These studies demonstrated the potential capability of using a computer in the detection of mammographic abnormalities. Gale et al. (17) and Getty et al. (18) have both reported on computer-based classifiers, which take as input diagnostically relevant features obtained from radiologists' readings of breast images. Getty et al. found that with the aid of the classifier, community radiologists performed as well as unaided expert mammographers in making benign vs. malignant decisions. Swett et al. (19) developed an expert system to provide visual and cognitive feedback to radiologists using a critiquing approach combined with an expert system.

At the University of Chicago, we have shown that the computerized analysis of mass lesions (21,23) and clustered microcalcifications (22,24) on digitized mammograms yields performances similar to an expert mammographer and significantly better than average radiologists in the task of distinguishing between malignant and benign lesions.

Ultrasound is a digital modality that is amenable to the application of computer-aided diagnosis techniques that ultimately could be used in a real-time fashion, at the time of examination, to improve diagnostic accuracy. Given that sonographic interpretation is a subjective process, and criteria have been developed that may allow differentiation of benign and malignant solid breast masses, it is reasonable to assume that CAD techniques applied to sonographic images would also improve radiologists' performance, particularly when combined with corresponding mammographic data (25). Recently, Garra et al. showed promising results in the use of computer-extracted features derived from co-occurrence matrices of images of breast lesions (26).

In this paper, we present methods for the computer analysis of breast lesions in gray-scale ultrasound images in order to discriminate between malignant and benign lesions.

Materials and Methods

Database

Masses were viewed sonographically by filming representative images in orthogonal planes. The ultrasound examinations were performed using an Ultramark 9 with High Definition Imaging (HDI) ultrasound machine from Advanced Technology Laboratories (ATL; Bothell, Washington, USA) with a high frequency, 7.5 Mhz, electronically focused, near-field imaging probe. The static images of the lesions that did not contain overlaid cursors or color Doppler signals were used for computer analysis.

The ultrasound films were retrospectively collected and then digitized by a laser film scanner (KFDR-S; Konica Corporation, Tokyo, Japan) with a scanner pixel size of 0.1 mm and 10-bit quantization. Each multiformat film contained only one ultrasound image. While film digitization is not the optimal approach to acquiring digital ultrasound data, it was the only one available for this initial study.

The ages of the 58 patients in this study ranged from 35 years to 89 years with a mean age of 53 years. They had 78 masses presented by 184 digitized ultrasound images. Benign lesions were confirmed by biopsy, cyst aspiration, or image interpretation alone while the malignant lesions were confirmed by biopsy. 144 images were from 43 patients with 62 benign lesions and 40 images from 15 patients with 16 malignant lesions. The benign lesions included simple cysts, complex cysts, and solid masses. Of the 62 benign lesions in the database, 19 (all solid lesions) lesions were proven by biopsy, 5 lesions (one solid lesion and 4 complex cysts) were proven by cyst aspiration, and 38 lesions (4 solid lesions and 38 cysts) were deemed benign by visual interpretation of the ultrasound images alone (e.g., simple cysts). All 16 malignant lesions were proven by biopsy.

The lesions in the database were further subcategorized into an "equivocal" category based on the necessity of performing an interventional procedure to determine their status. The 24 benign lesions that were proven via either biopsy or cyst aspiration, and the 16 malignant lesions, made up the equivocal database for a total of 40 lesions. The purpose of this subcategorization was to determine the ability of the computer features to distinguish between benign and malignant lesions that required an interventional procedure (i.e., cyst aspiration or biopsy) for definitive diagnosis.

Manual lesion segmentation and ROI selection

Once digitized, the ultrasound images were displayed on a CRT monitor and a breast imaging radiologist (DEW) outlined the approximate margins of each lesion. Figures 1 (a-d) show ultrasound images of breast lesions with margins outlined. Regions-of-interest (ROIs), 32 x 32 pixels in size, were selected from regions within and posterior to the lesions. Features were calculated based either on the manually-extracted lesion margin or on the 32 x 32 ROI.

Automated feature extraction

Four main types of lesion characteristics were investigated, including the margin of the lesion, the shape of the lesion, the lesion homogeneity (texture), and the nature of the posterior acoustic attenuation pattern. Table 1 lists these characteristics along with their relationships to benign and malignant lesions.

The lesion characteristics were quantified using various computer-extracted features. Table 2 lists the computer-extracted features used in distinguishing between malignant and benign lesions. Features were either calculated along or within the lesion margin, or within a 32 x 32 pixel ROI (placed within the central portion of the lesions or posterior to the lesion).

To quantify lesion margin characteristics, gradient analysis was performed along a computer-expanded margin of the lesion. In this analysis, first the manually-extracted margin is expanded using morphological filtering. Then, this region is processed by a Sobel filter in order to obtain the gradient and its direction at each pixel. The normalized radial gradient is calculated to quantify the margin sharpness as well as the degree of margin irregularity (shape) (21,27). The normalized radial gradient (22,27) is given by

$$\text{Normalized radial gradient} = \frac{\sum_{P \in margin} \cos\varphi \sqrt{D_x^2 + D_y^2}}{\sum_{P \in margin} \sqrt{D_x^2 + D_y^2}}$$

where D_x is the gradient in the x-direction, D_y is the gradient in the y-direction, and φ is the angle between the gradient vector and the radial gradient. A lower value of the normalized radial gradient indicates a less distinct margin.

The geometric measure of shape in terms of a short-to-long axis ratio of each lesion was determined using the image data along the margin. Note here that the short-to-long axis ratio corresponds to a depth-to-width ratio in order to extract out the orientation of the long axis. Cysts tend to be elipsoid resulting in a depth-to-width ratio of much less than one, whereas malignant lesions tend to have a vertical or round axial orientation (28).

Texture can be described by spatial relationships between image pixels using changes in the intensity patterns and gray levels. Texture characteristics of the homogeneity within the lesion were determined using a measure of coarseness (29). The texture measure of coarseness (local uniformity) is given by (29):

$$COS = [\sum_{i}^{G_h} p_i s(i)]^{-1}$$

where G_h is the highest gray-level value in the ROI, p_i is the probability of occurrence of gray-level value i, N is the width of the ROI (N=32), d is the neighboring size (half of the operating kernel size,W), and the ith entry of s is given by

$$s(i) = \begin{cases} \sum |i - A_i| & \text{for } i \in \{N_i\} \text{ if } N_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$\{N_i\}$ is the set of pixels having gray level i

$$A_i = \frac{1}{W-1} \sum_{p=-d}^{d} \sum_{p=-d}^{d} f(x+p, y+q) \qquad (p,\ q) \neq (0,0) \text{ to exclude } (x,y)$$

$$W = (2d+1)^2 \qquad (d=3)$$

A lower value of coarseness corresponds to a finer visual texture.

The computerized assessment of posterior acoustic attenuation or enhancement associated with different lesions was determined in two ways: (1) by comparing the gray level values within the lesion with those posterior to the lesion and (2) by comparing the gray levels posterior to the lesion with those in the adjacent tissue at the same depth. This was performed to quantify the amount of any posterior acoustic shadowing or enhancement. For example, benign lesions are often associated with posterior enhancement while malignant lesions may produce posterior shadowing. Simple cysts that are anechoic will cause less attenuation of the ultrasound waves than surrounding parenchyma, and thereby cause relative hyperechogenicity posterior to the lesion. In this analysis, ROIs, 32 x 32 pixels in size, were placed within the lesion, posterior to the lesion, and in the adjacent tissue at the same depth; and the differences in average gray levels was calculated to quantify posterior acoustic attenuation.

The feature value for a given lesion was obtained by averaging that feature value over all views of the lesion, since each lesion had two to five views from one clinical examination.

Linear discriminant analysis (LDA) was used to merge the four individual, computer-extracted features into a single index related to the estimate of the likelihood of malignancy. In linear discriminant analysis, the discriminant function is formulated by a linear combination of the individual features (30). Both consistency and round-robin runs were performed. In round-robin analysis, the discriminant function is trained on all but one case and then tested on the left-out case. This is repeated until all cases have been individually tested.

Evaluation

Receiver operating characteristic (ROC) analysis (31) was used to evaluate (by case; not by image) the performance of individual features in the task of distinguishing benign from malignant lesions. The decision variable for the ROC analysis was each individual feature. The area under the ROC curve (A_Z) was used as an indicator of performance. Specificity at high sensitivity is very clinically relevant since the "cost" of missing a cancer is greater than performing an interventional procedure for a benign lesion. Therefore, in addition, we calculated the performance of the features in the high sensitivity range (true-positive fraction $TPF_0 > 0.90$) using the partial area index ,TPF_0A_Z', which is the portion of the area under the ROC curve that lies above TPF_0 divided by the constant ($1-TPF_0$) (32). Both A_Z and partial A_Z values were calculated for the entire database as well as for the "equivocal" database.

Results

The A_Z values for the various computer-extracted ultrasound features ranged from 0.54 to 0.85 in the task of distinguishing benign from malignant lesions. These are given in Table 2 for both the entire database and the equivocal database. Since missing a cancer is more clinically significant than performing an interventional procedure for a benign lesion, we have used the

partial area index to quantify the performance of the features at a high-sensitivity level (32). These TPF_0A_Z' values are also given in Table 2 for the entire database and the equivocal database.

Using the first four features, linear discriminant analysis consistency runs yielded A_Z values of 0.95 and 0.93 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively, and the round-robin runs yielded A_Z values of 0.94 and 0.87 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively. (Table 2) When the second posterior acoustic attenuation feature was used, linear discriminant analysis consistency runs yielded A_Z values of 0.94 and 0.93 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively, and the round-robin runs yielded A_Z values of 0.92 and 0.86 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively. (Table 2)

Discussion

Figures 2 (a and b) show cluster plots of the coarseness and the margin feature for malignant and benign lesions for the entire database and the equivocal database, respectively. Figures 3 (a and b) show cluster plots of the depth-to-width ratio and the (first) posterior acoustic attenuation feature for malignant and benign lesions for the entire database and the equivocal database, respectively. It is apparent that malignant lesions tend to exhibit less distinct margins and more posterior shadowing than do the benign lesions as decumented by visual ultrasound criterion (10,12,26,28). It is interesting to note that many benign lesions not in the equivocal database had very "fine" texture (i.e., a low coarseness feature) since many were cysts and anechoic, whereas the benign solid leisons tended to have coarse texture.

Figures 4 (a and b) show the performance in terms of ROC curves for the features characterizing margin and acoustic attenuation, respectively, in the task of distinguishing between malignant and benign lesions for the entire database and the equivocal database. It should be noted that the ROC curves are lower for the equivocal database than for the entire database indicating that, as for radiologists, the benign lesions in the equivocal database are more difficult to distinguish from malignant lesions. This is further illustrated in Figures 5 a-c which show the histograms of the feature values (for the malignant lesions, the benign lesions in the equivocal database, and the remaining benign lesions) for margin, texture, and posterior acoustic attenuation, respectively. It is apparent that malignant lesions tend to exhibit a coarse texture, less distinct margins, and posterior shadowing. However, it should also be noted that substantial overlap of the features exist for the malignant lesions and the benign lesions in the equivocal database.

Table 2 and Figure 6 show the perfomance of the linear discriminant function in the task of distinguishing between malignant and benign lesions in both the entire and equivocal database. Of note is the partial A_Z values and the shape of the ROC curves. Extraction from the fitted ROC curves, for the equivocal database in which all lesions had had a clinical procedure of cyst aspiration or biopsy, indicates that at a high sensitivity level (90%) for malignant cases, 30% of the benign cases were classified as benign and thus, could potentially have avoided biopsy. It is apparent that the combined use of the four computer-extracted features yield superior performance.

While film digitization was not the optimal approach to acquiring our digital ultrasound database, it was the only one available for this initial study. It should be noted that even with the limited image quality obtained from the film digitization of the multi-format film ultrasound images, we were able to observe computer features capable of distinguishing between malignant

and benign lesions. Computerized analysis of direct digital ultrasound data is expected to yield an improvement in the discriminating ability of the various features, especially for texture features of the lesion interior. Our future database collection will include direct digital data.

The purpose of this study was to determine if computer-extracted features from ultrasound images of breast lesions have the potential to discriminate between malignant and benign lesions, and thus ultimately aid in the reduction of unnecessary biopsies. This potential has been shown even though the number of lesions in the database is small. In addition, the features chosen agree well with those used by radiologists when interpreting breast ultrasound images. It should be noted that the computer-extracted features were obtained from radiologist-drawn margins. In the future, the subjectiveness of human-drawn margins will be eliminated with the use of computer segmentation.

Ultrasound images of the breast can yield information on the interior of the lesion (homogeneity) as well as the interface between the lesion and its surround. This is the reason that ultrasound is used to distinguish between solid and cystic lesions. Gradient analysis of the margin yields information on the lesion margin including sharpness and shape. Geometric features relating the depth-to-width ratio of the lesion is useful in that although many solid lesions may be ellipsoid, the orientation of the ellipse with regard to the skin is important in distinguishing between benign and malignant lesions. Computerized analysis also allows for the objective assessment of posterior acoustic shadowing and enhancement which can also aid in distinguishing between benign and malignant lesions. It should be noted though that because shadowing depends on gain settings, scanning parameters should be set more automatically than they are at present.

Use of linear discriminant analysis as a classifier with which to merge the four computer-extracted features proved useful in improving performance in the task of distinguishing between

malignant and benign lesions in both the entire and equivocal database. As the database increases, other classifiers such as artificial neural networks will be intevestigated as means with which to merge the features into an estimate of the likelihood of malignancy.

Conclusion

We are developing methods for the computer analysis of breast lesions in ultrasound images. In our study, we automatically extracted and calculated features related to lesion margin, lesion shape, texture (homogeneity) within the lesion, and posterior acoustic attenuation. From ROC analysis of the computer-extracted features, the features based on margin characteristics yielded A_Z values of 0.85 and 0.75, in the task of distinguishing between benign and malignant lesions in the entire database and in an "equivocal" database, respectively. The equivocal database consisted of lesions that had been proven to be benign or malignant by either cyst aspiration or biopsy. Linear discriminant analysis consistency runs yielded A_Z values of 0.95 and 0.93 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively, and the round-robin runs yielded A_Z values of 0.94 and 0.87 in the task of distinguishing between benign and malignant lesions in the entire database and the equivocal database, respectively. Our results indicate that the computerized analysis of ultrasound images has the potential to increase the specificity of breast sonography. These promising results warrant further development and testing on a large direct-digital database.

Acknowledgements

The authors would like to thank Dr. Fred Winsberg for contributing to the database. The authors would also like to thank Dr. Ulrich Bick, Jason Rubenstein, Michael Chinander, and Dr. Samuel Armato, III, for useful discussions. This research was supported in parts by NIH grant P20 CA66132 and U.S. Army Medical Research and Materiel Command, DAMD17-98-1-8194

Maryellen L. Giger is a shareholder in R2 Technology, Inc. (Los Altos, CA). It is the University of Chicago Conflict of Interest Policy that investigators disclose publicly actual or potential significant financial interest which would reasonably appear to be directly and significantly affected by the research activities.

References

1. American Cancer Society: Cancer Facts and Figures -- 1998. New York, NY, 1998.
2. Feig SA: Decreased breast cancer mortality through mammographic screening: Results of clinical trials. Radiology 167:659-665, 1988.
3. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O: Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 30:187-210, 1992.
4. Smart CR, Hendrick RE, Rutledge JH, Smith RA: Benefit of mammography screening in women ages 40 to 49 years: Current evidence from randomized controlled trials. Cancer 75:1619-26, 1995.
5. Bassett LW, Jackson VP, Jahan R, Yao SF, Gold RH: Diagnosis of Diseases of the Breast Philadelphia: W. B. Saunders, 1997.
6. Kopans DB: Breast Imaging. Philadelphia: JB Lippincott, 1998.
7. Brown ML, Houn F, Sickles EA, Kessler LG: Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnostic procedures. AJR 165:1373-1377, 1995.
8. Jackson VP: The role of US in breast imaging. Radiology 177:305-311, 1990.
9. Kolb TM, Lichy J, Newhouse JH: Occult cancer in women with dense breasts: Detection with screening US -- Diagnostic yield and tumor characteristics. Radiology 207: 191-199, 1998.
10. Tohno E, Cosgrove DO, Sloane JP: Ultrasound Diagnosis of Breast Diseases. Churchill Livingstone, Edinburgh, 1994, pp. 50-73.
11. Fornage BD, Lorigan JG, Andry E: Fibroadenoma of the breast: sonographic appearance. Radiology 172:671-675, 1989.
12. Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA: Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 196:123-134, 1995.
13. Huber S, Delorme S, Knopp MV, Junkermann H, Zuna I, von Fournier D, van Kaick G: Breast tumors: computer-assisted quantitative assessment with color Doppler US. Radiology 192:797-801, 1994.
14. Birdwell RL, Ikeda DM, Jeffrey SS, Jeffrey RB, Jr.: Preliminary experience with power Doppler imaging of solid breat masses. AJR 169: 703-707, 1997.
15. Giger ML: Computer-aided diagnosis. In: Syllabus: A Categorical Course on the Technical Aspects of Breast Imaging, edited by Haus A, Yaffe M. Oak Brook, IL: RSNA Publications, 1993, pp. 272-298.
16. Vyborny CJ, Giger ML: Computer vision and artificial intelligence in mammography. AJR 162:699-708, 1994.
17. Gale AG, Roebuck EJ, Riley P, Worthington BS, et al.: Computer aids to mammographic diagnosis. British Journal of Radiology 60: 887-891, 1987.
18. Getty DJ, Pickett RM, D'Orsi CJ, Swets JA: Enhanced interpretation of diagnostic images. Invest. Radiol. 23: 240-252, 1988.
19. Swett HA, Miller PA: ICON: A computer-based approach to differential diagnosis in radiology. Radiology 163: 555-558, 1987.
20. Giger ML, Vyborny CJ, Schmidt RA: Computerized characterization of mammographic masses: Analysis of spiculation. Cancer Letters 77: 201-211, 1994.

21. Huo Z, Giger ML, Vyborny CJ, Bick U, Lu P, Wolverton DE, Schmidt RA: Analysis of spiculation in the computerized classification of mammographic masses" Medical Physics 22:1569-1579, 1995.
22. Jiang Y, Nishikawa RM, Wolverton DE, Metz CE, Giger ML, Schmidt RA, Vyborny CJ, Doi K: Automated feature analysis and classification of malignant and benign clustered microcalcifications. Radiology 198:671-678, 1996.
23. Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Schmidt RA, Doi K: Automated computerized classification of malignant and benign mass lesions on digitized mammograms. Academic Radiology 5: 155-168, 1998.
24. Jiang Y, Nishikawa, et al.: Improving breast cancer diagnosis with computer-aided diagnosis. Academic Radiology, (in press), 1998.
25. Giger ML, Huo Z, Wolverton DE, Vyborny C, Moran C, et al.: Computer-aided diagnosis of digital mammographic and ultrasound images of breast mass lesions. In Digital Mammography 1998, Eds: Karssemeijer N, Thijssen M, Hendriks J, van Erning L, Kluwer Academic Publishers, Dordrecht, The Netherlands, pps: 143-147, 1998.
26. Garra BS, Krasner BH, Horii SC, Ascher S, Mun SK, Zeman RK: Improving the distinction between benign and malignant breast lesions: The value of sonographic texture analysis. Ultrasonic Imaging, 15: 267-285, 1993.
27. Bick U, Giger ML, et al: A new single-image method for computer-aided detection of small mammographic masses. Proc. CAR '95, LemkeHU, Inamura K, Jaffe CC, Vannier MW, eds. pgs. 357-363, 1995.
28. Sohn C, Blohmer J-U, Hamper UM. Breast Ultrasound: A Systematic Approach to Technique and Image Interpretation, Thieme, New York, 1998.
29. Amadasun M, King R: Textural features corresponding to textural properties. IEEE Trans. on Syst., Man, & Cybernetics 19:1264-1274, 1989.
30. Lachenbruch PL. Discriminant analysis. London, Hafner Press, 1975.
31. Metz CE: Some practical issues of experimental design and data analysis in radiologicalROC studies. Invest. Radiol. 24: 234-245, 1989.
32. Jiang Y, Metz CE, Nishikawa RM: A receiver operating characteristics partial area index for highly sensitive diagnostic tests. Radiology 201: 745-750, 1996.

Captions

Figure 1 Examples of (a) a simple cyst, (b) a complex cyst, (c) a benign solid lesion, and (d) a malignant lesion along with the radiologist-drawn lesion margin.

Figure 2 Cluster plots indicating feature values for margin and texture features for (a) the entire database and (b) the "equivocal" database.

Figure 3 Cluster plots indicating feature values for lesion shape and posterior acoustic attenuation for (a) the entire database and (b) the "equivocal" database.

Figure 4 ROC curves illustrating the performance of the computer-extracted features of (a) margin and (b) posterior acoustic attenuation in the task of distinguishing between malignant and benign lesions for both the entire database and the "equivocal" database.

Figure 5 Histograms of the computer-extracted features of (a) margin, (b) texture, and (c) posterior acoustic attenuation showing the distribution for the benign non-equivocal database, the benign equivocal database, and the malignant database.

Figure 6 ROC curves illustrating the performance of the discriminant scores in the task of distinguishing between malignant and benign lesions for both the entire database and the "equivocal" database. Results are from round-robin analysis.

Table 1. Lesion characteristics investigated in this study.

	Benign (cystic or solid)	Malignant
Lesion Margin	smooth borders	angular margins spiculation
Lesion Shape	ellipsoid mildly lobulated	irregular depth:width > 0.8
Texture within Lesion	anechoic hyperechoic reverberation artifacts	hypoechoic
Posterior Acoustic Attenuation	posterior enhancement	posterior shadowing

Table 2. Computer-extracted features investigated in this study. Performance is given in terms of A_Z value and partial A_Z for the entire database and the "equivocal" database in the task of distinguishing between malignant and benign lesions.

Analysis	Region for Database	Entire Database $N_{case} = 78$		"Equivocal" $N_{case} = 40$	
		A_Z	$0.90A_Z'$	A_Z	$0.90A_Z'$
Lesion margin					
(1) normalized radial gradient	margin	0.85	0.46	0.75	0.28
Shape					
(2) depth-to-width ratio	margin	0.67	0.20	0.75	0.29
Texture within the lesion					
(3) coarseness	ROI	0.54	0.12	0.67	0.14
Posterior Acoustic Attenuation					
(4) difference in gray level between within lesion and posterior to lesion	ROIs	0.77	0.29	0.72	0.27
(5) Difference in gray level between posterior to lesion and adjacent tissue at same depth	ROIs	0.83	0.35	0.72	0.17
Linear Discriminant Analysis (features 1,2,3,4)					
consistency analysis		0.95	0.78	0.93	0.70
round-robin analysis		0.94	0.76	0.87	0.56
Linear Discriminant Analysis (features 1,2,3,5)					
consistency analysis		0.94	0.68	0.93	0.58
round-robin analysis		0.92	0.59	0.86	0.38

Simple Cyst

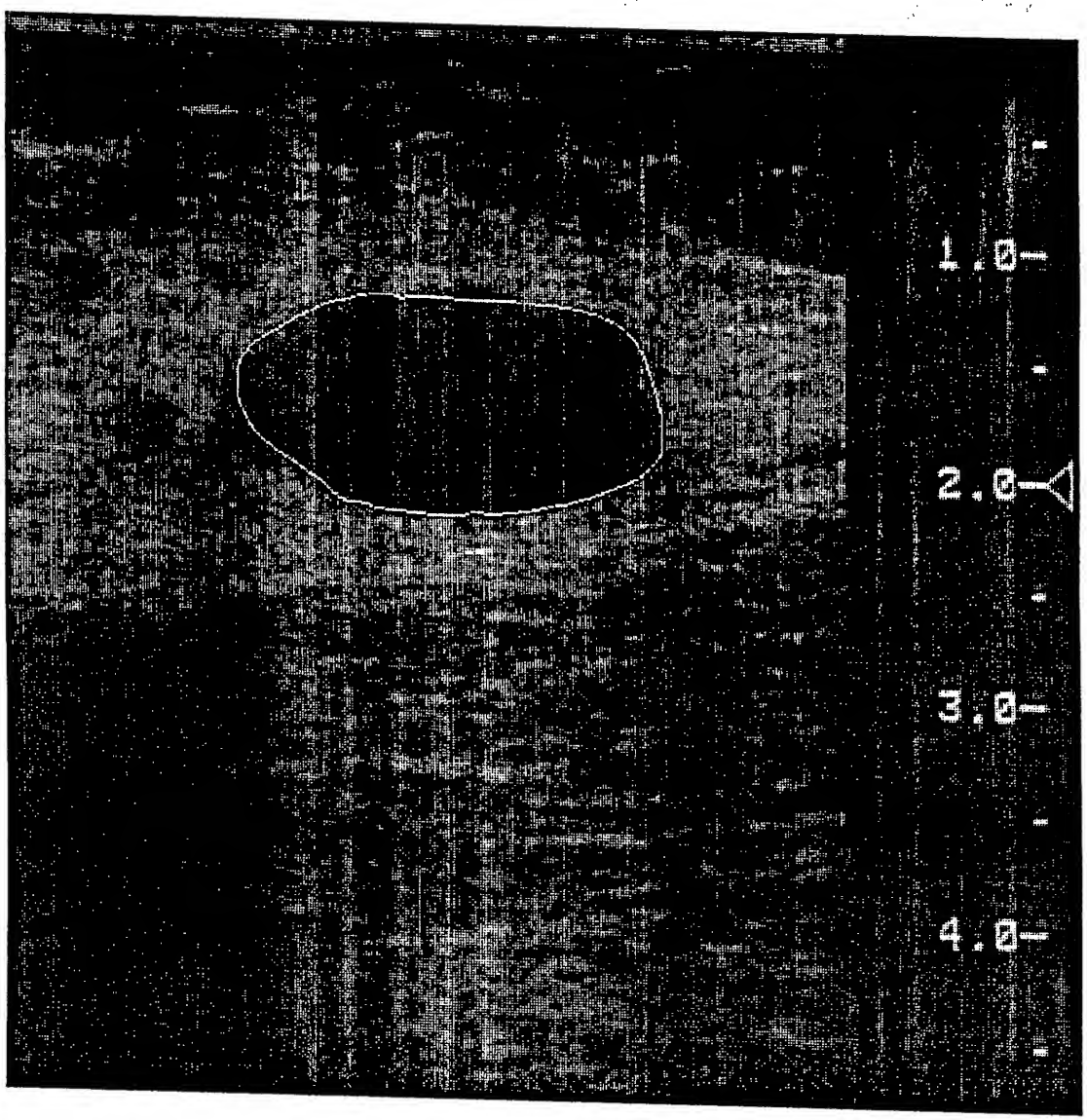


Fig 1 (a)

Fig #(6)

Complex Cyst

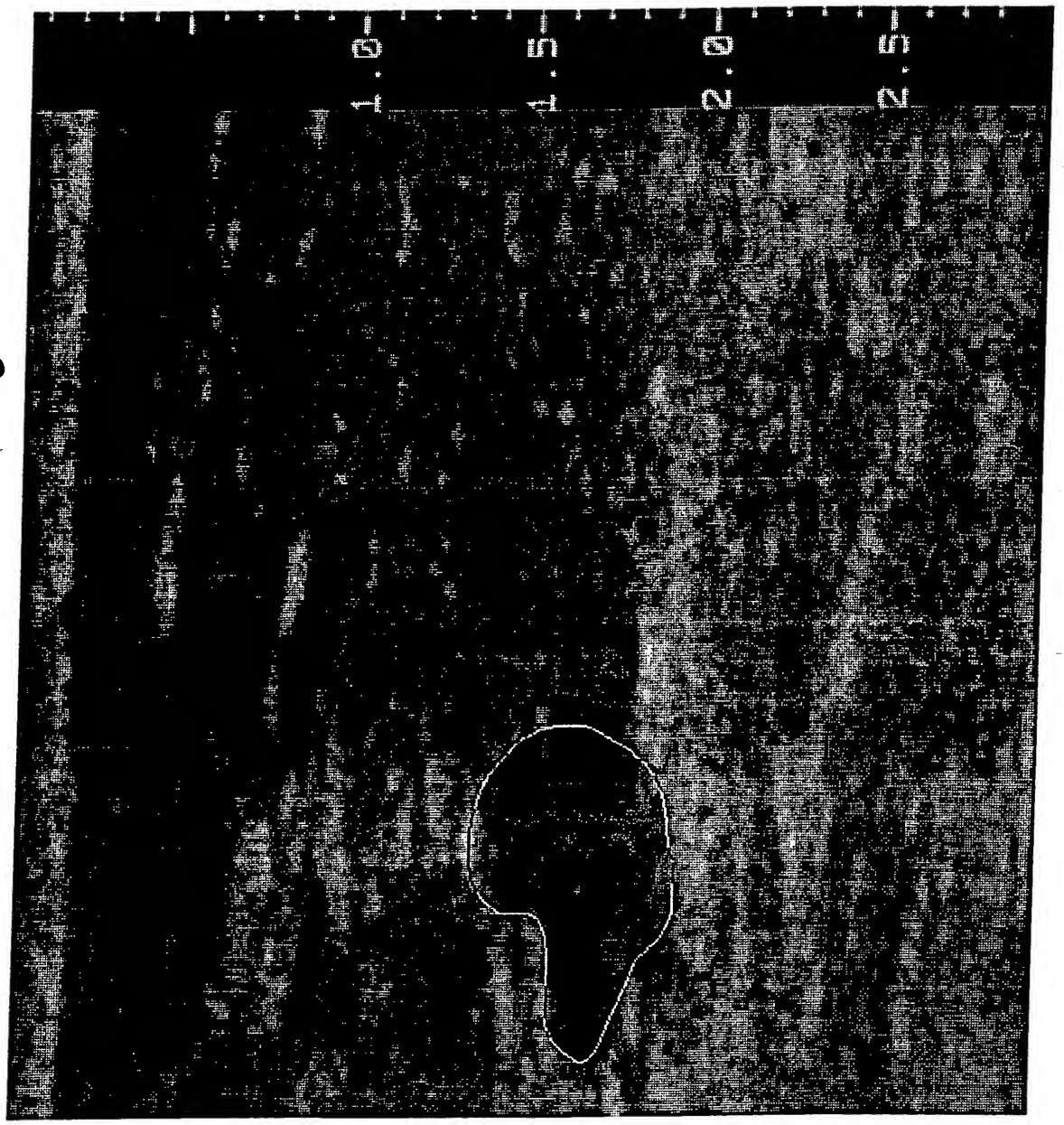
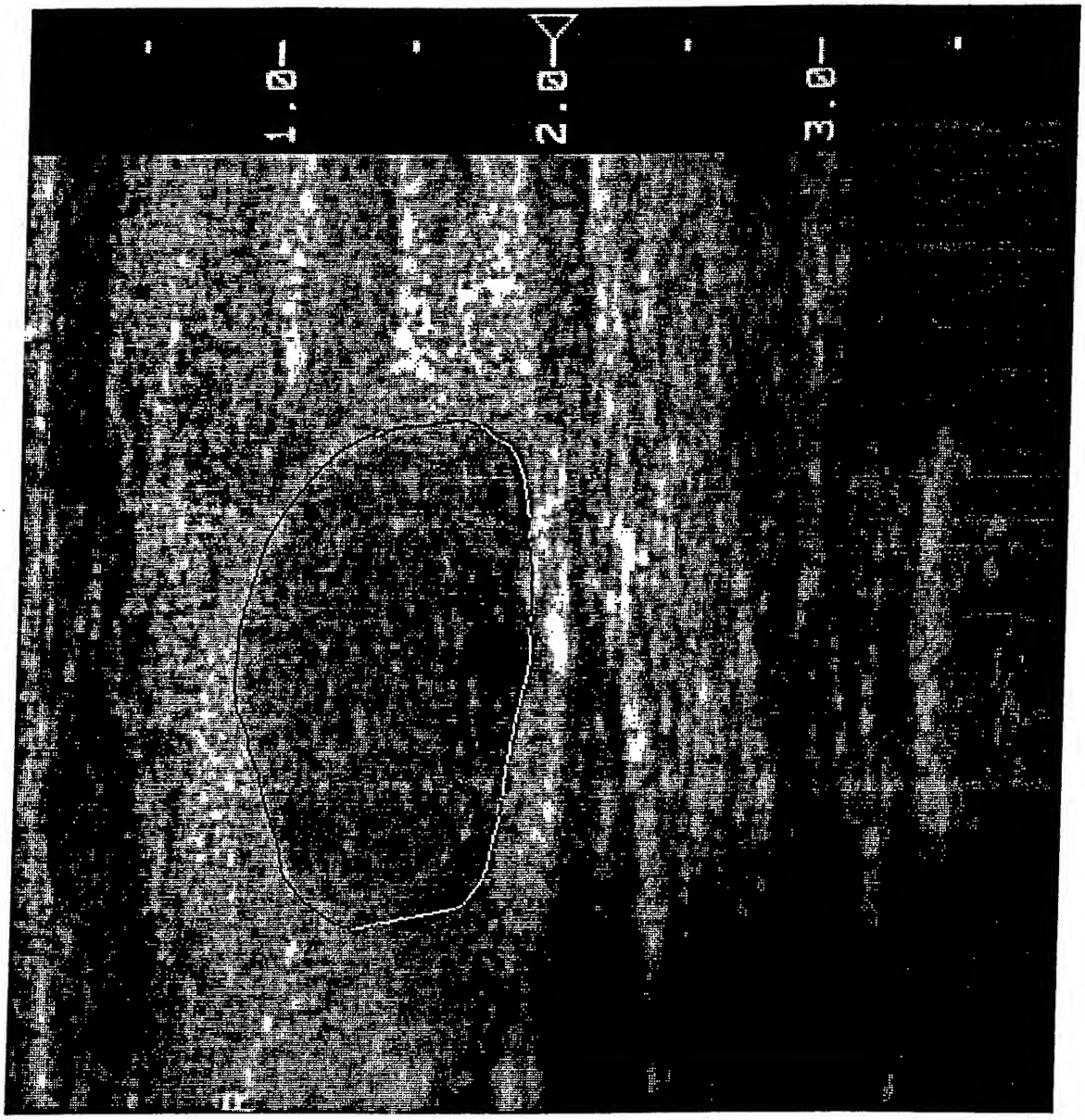


Fig 4(c)

Solid Lesion



Malignant Lesion

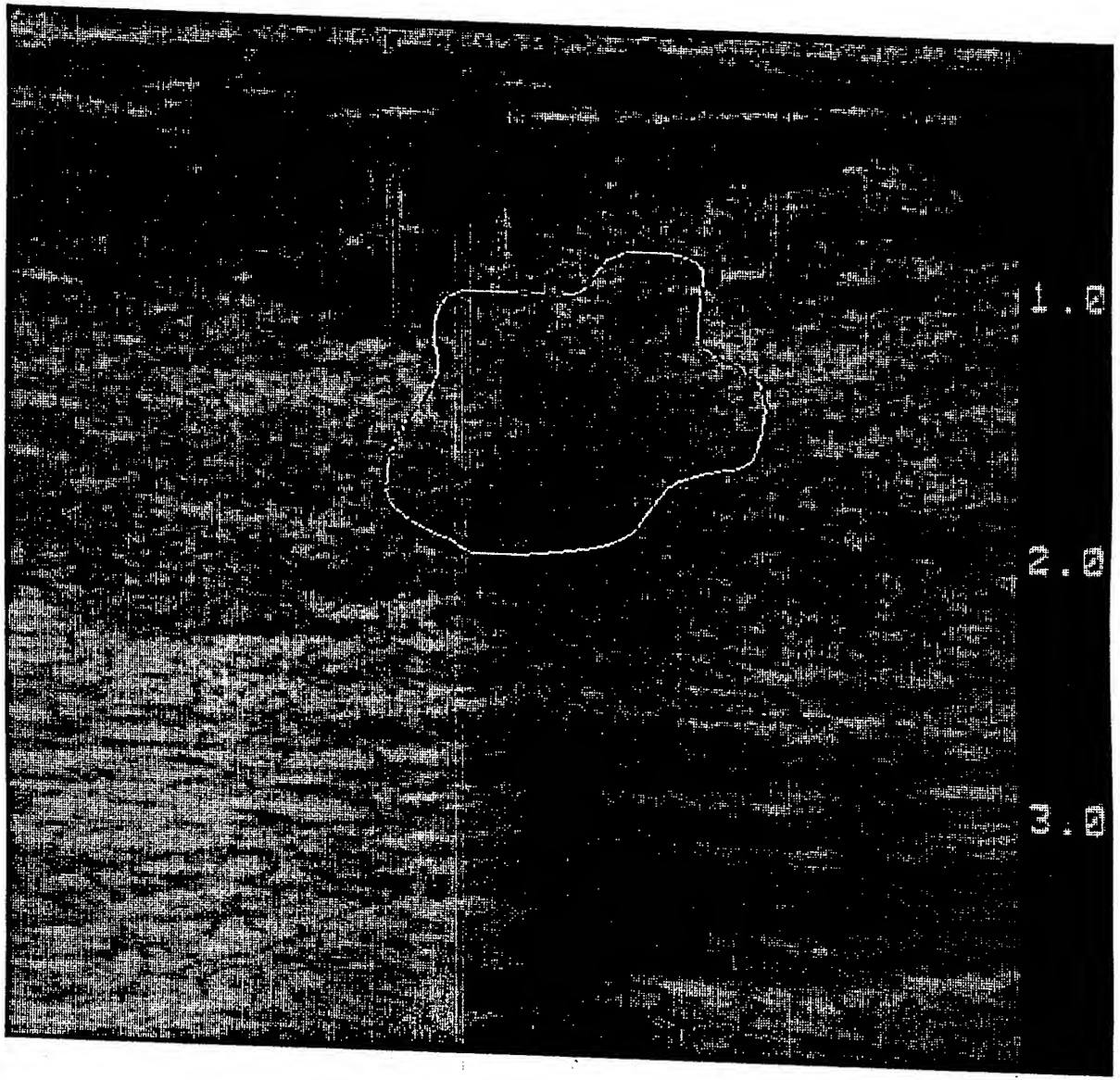
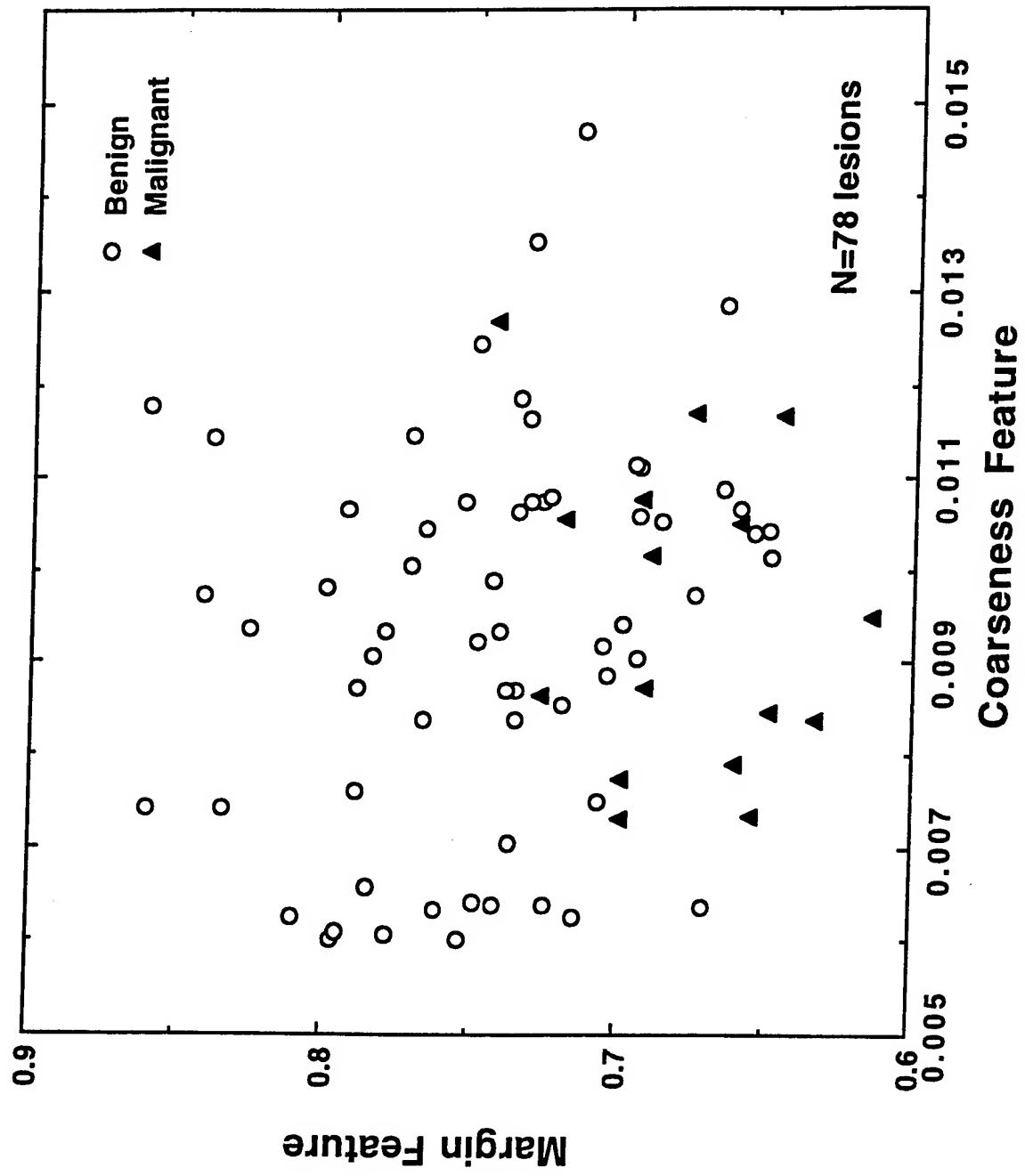
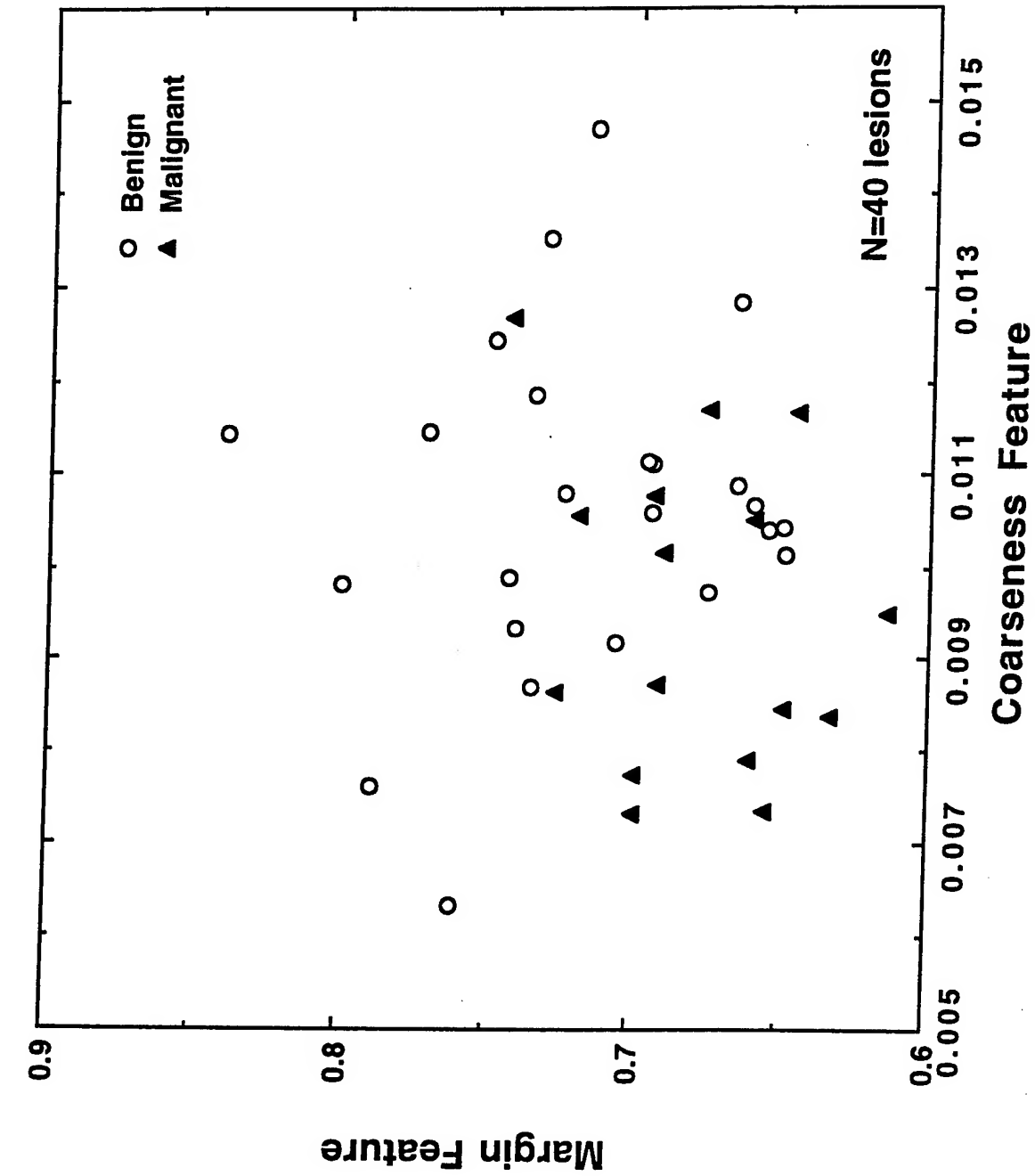
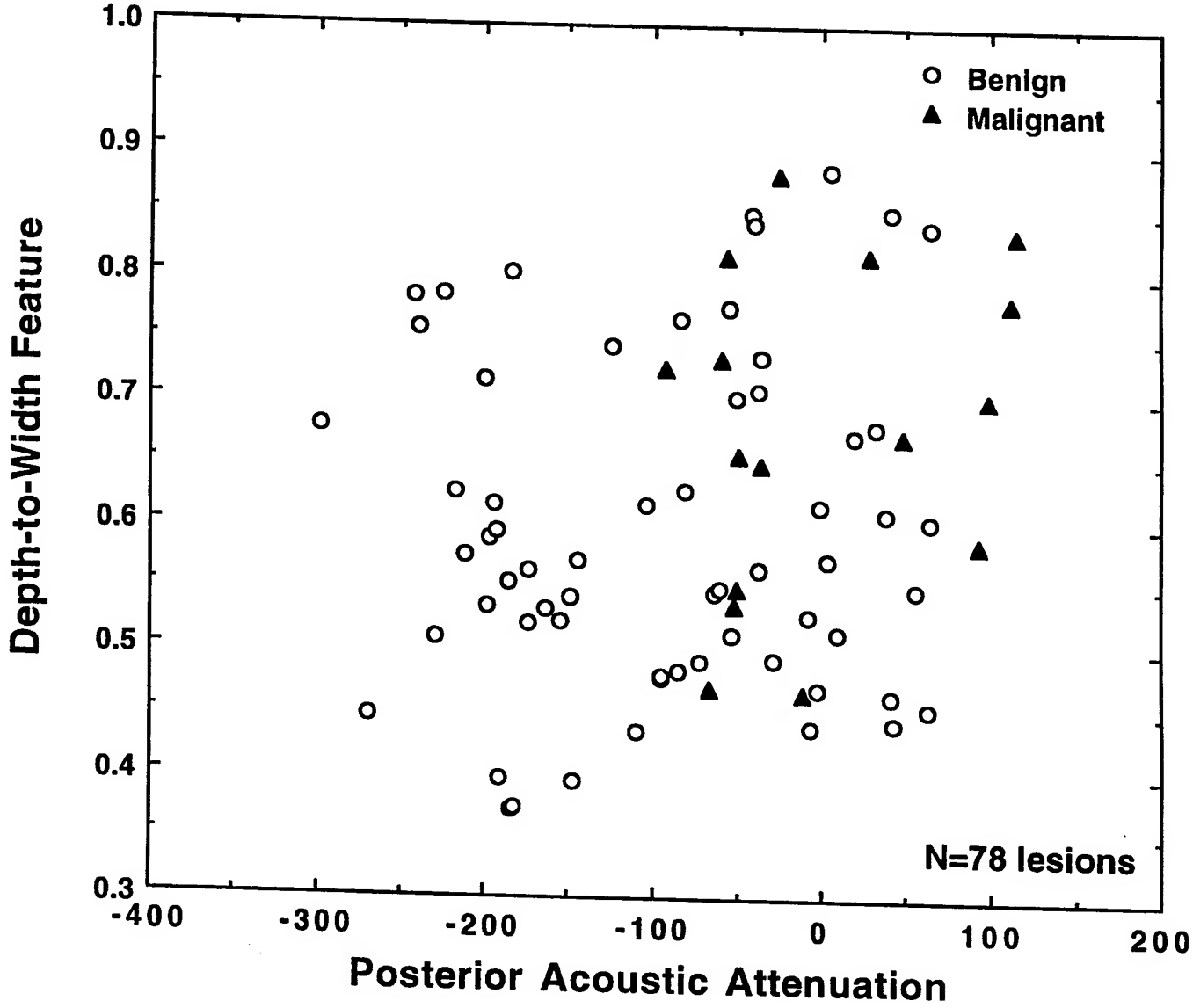
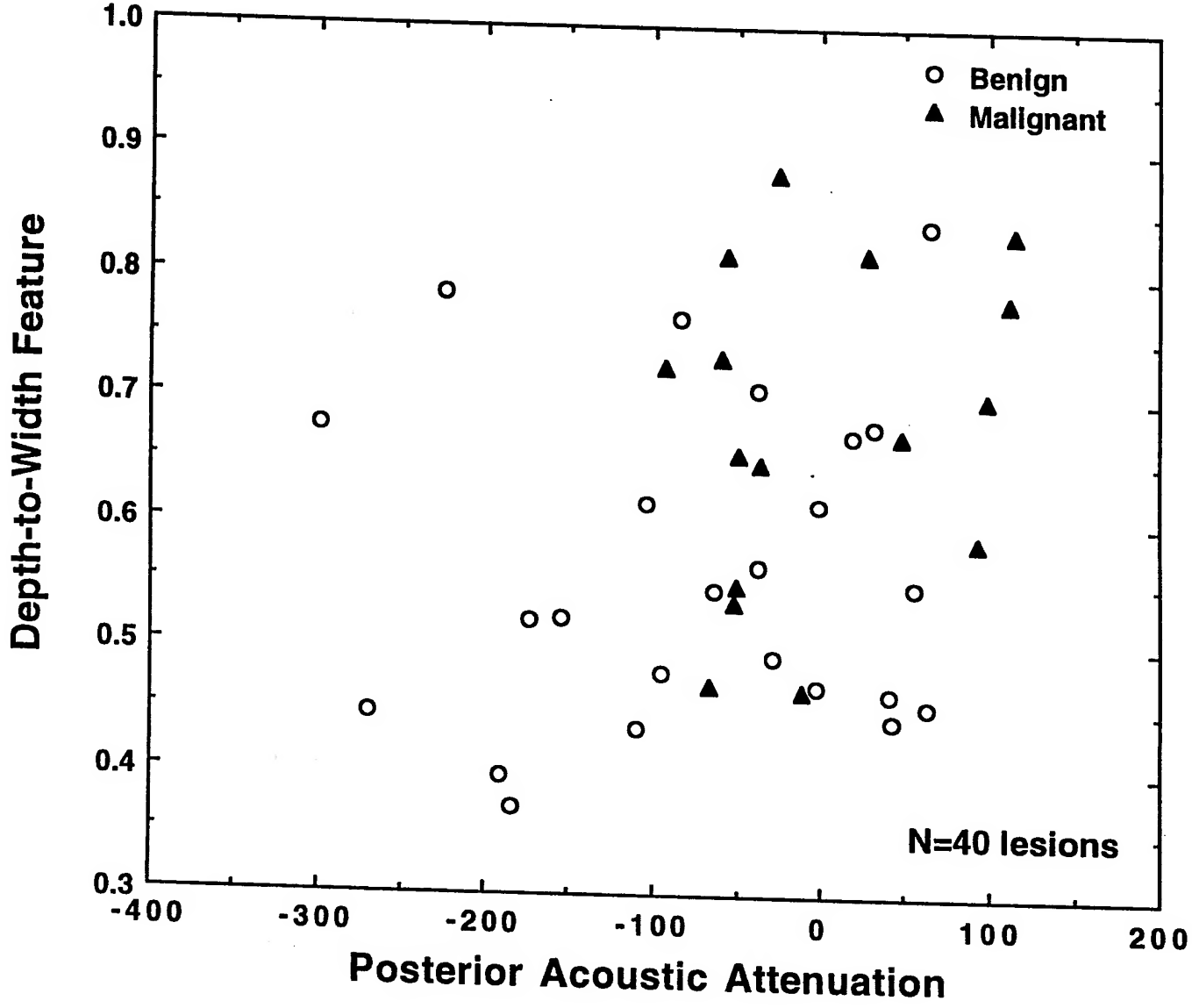


Fig 1(d)


Benign
Malignant
N=78 lesions
Margin Feature
Coarseness Feature
0.9
0.8
0.7
0.6
0.005
0.007
0.009
0.011
0.013
0.015


Benign
Malignant
N=40 lesions
Margin Feature
Coarseness Feature
0.9
0.8
0.7
0.6
0.005
0.007
0.009
0.011
0.013
0.015


Depth-to-Width Feature
Posterior Acoustic Attenuation
Benign
Malignant
N=78 lesions
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
-400
-300
-200
-100
0
100
200


Benign
Malignant
N=40 lesions
Depth-to-Width Feature
Posterior Acoustic Attenuation
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
-400
-300
-200
-100
0
100
200

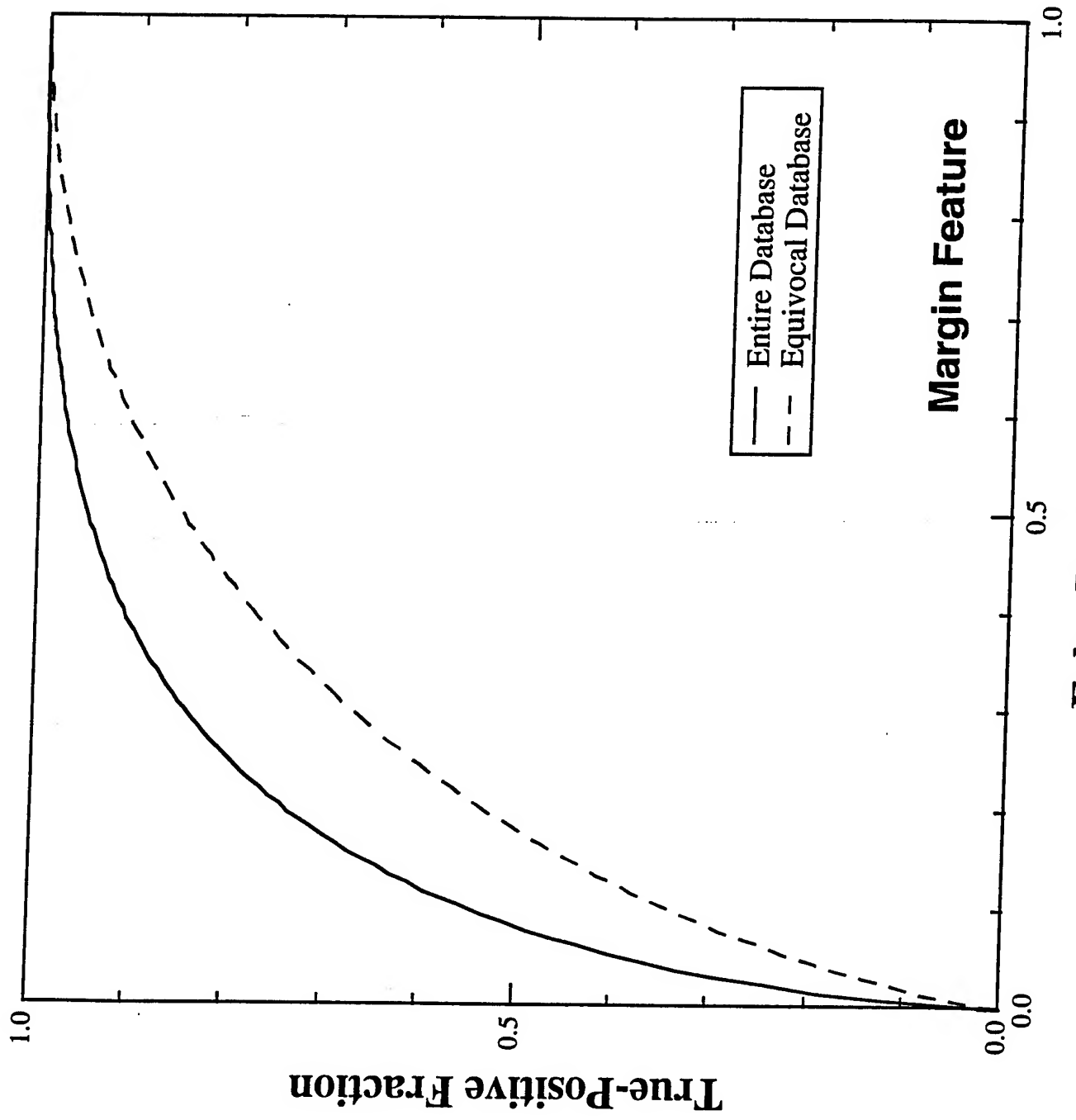
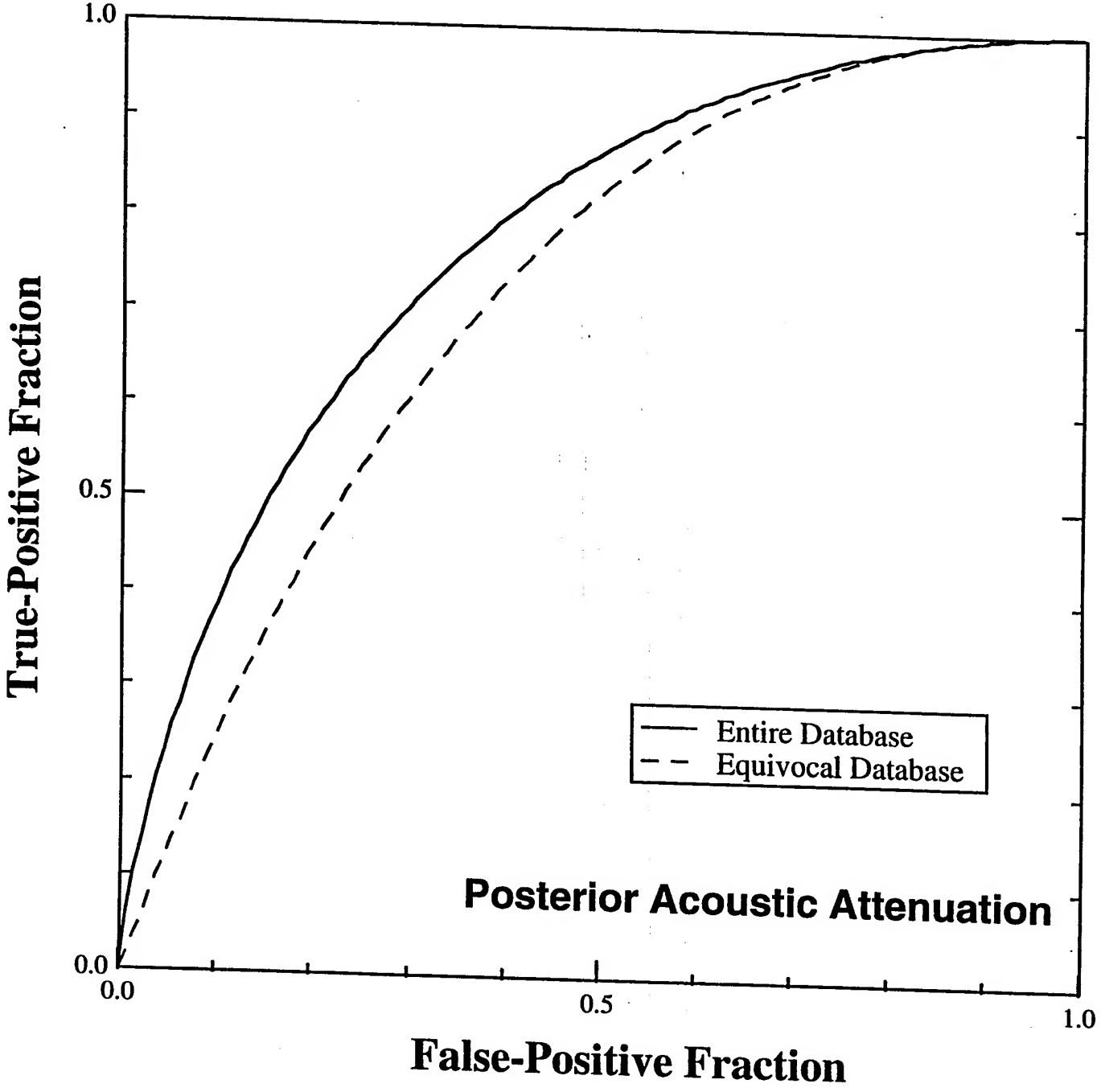
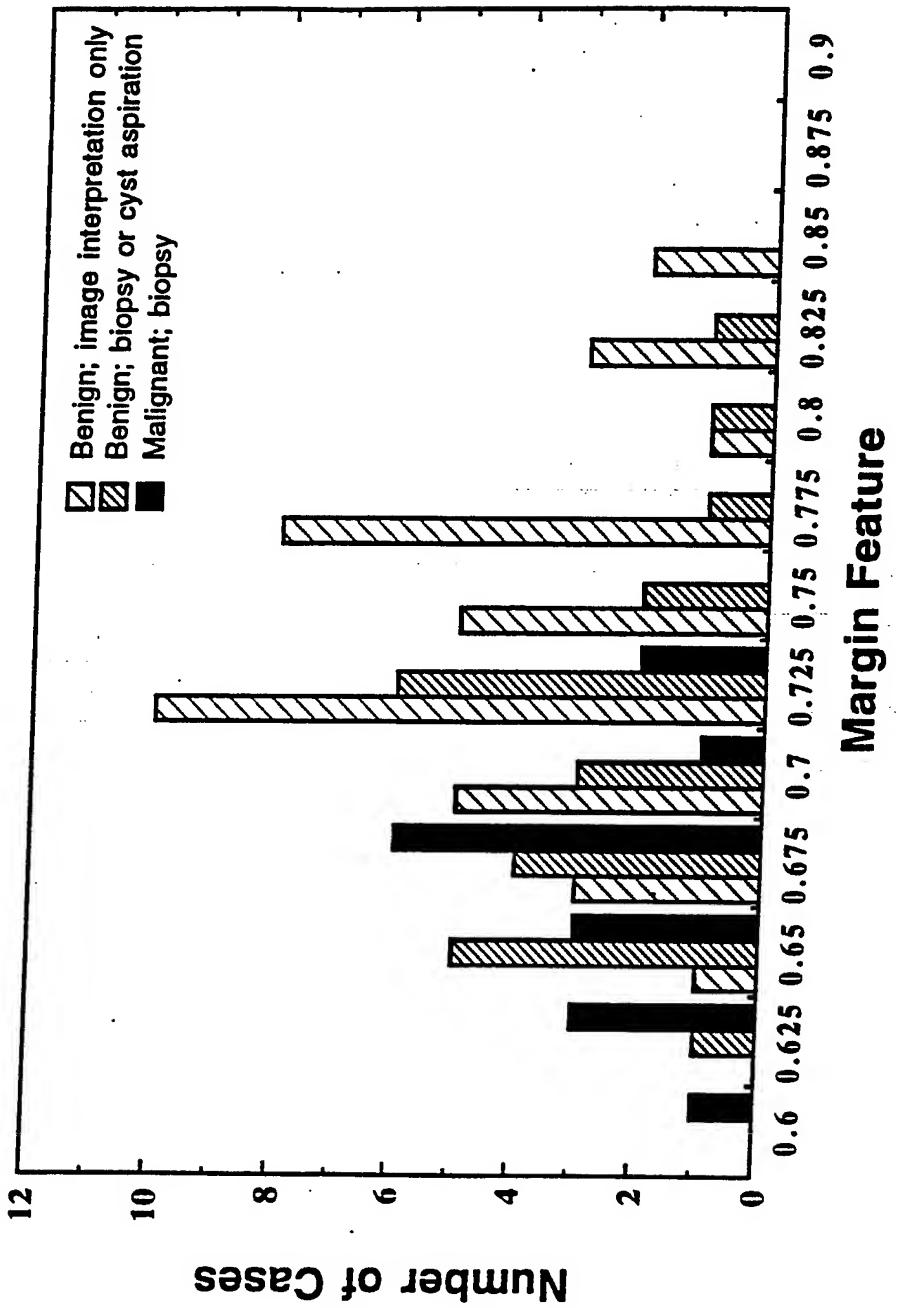
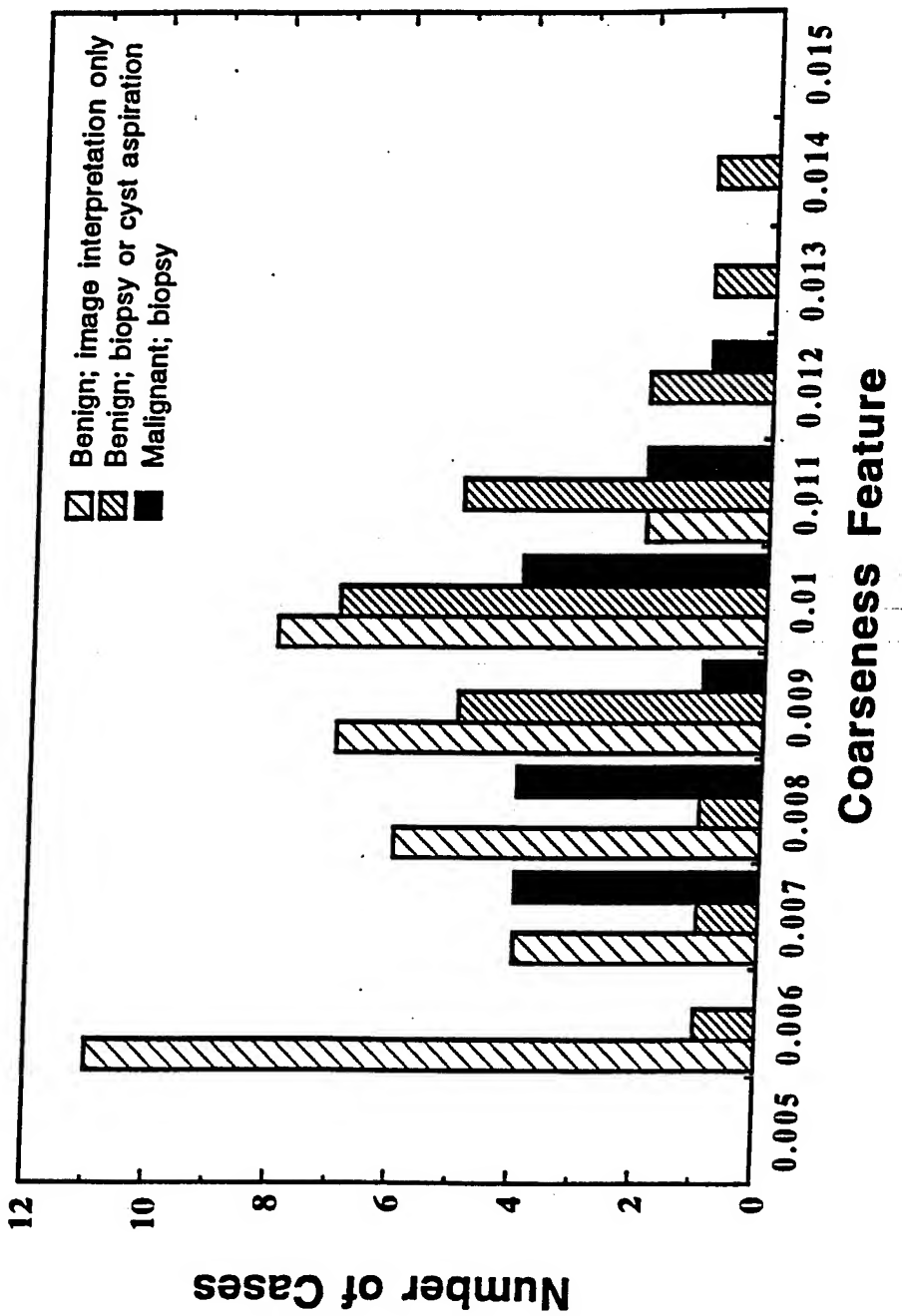
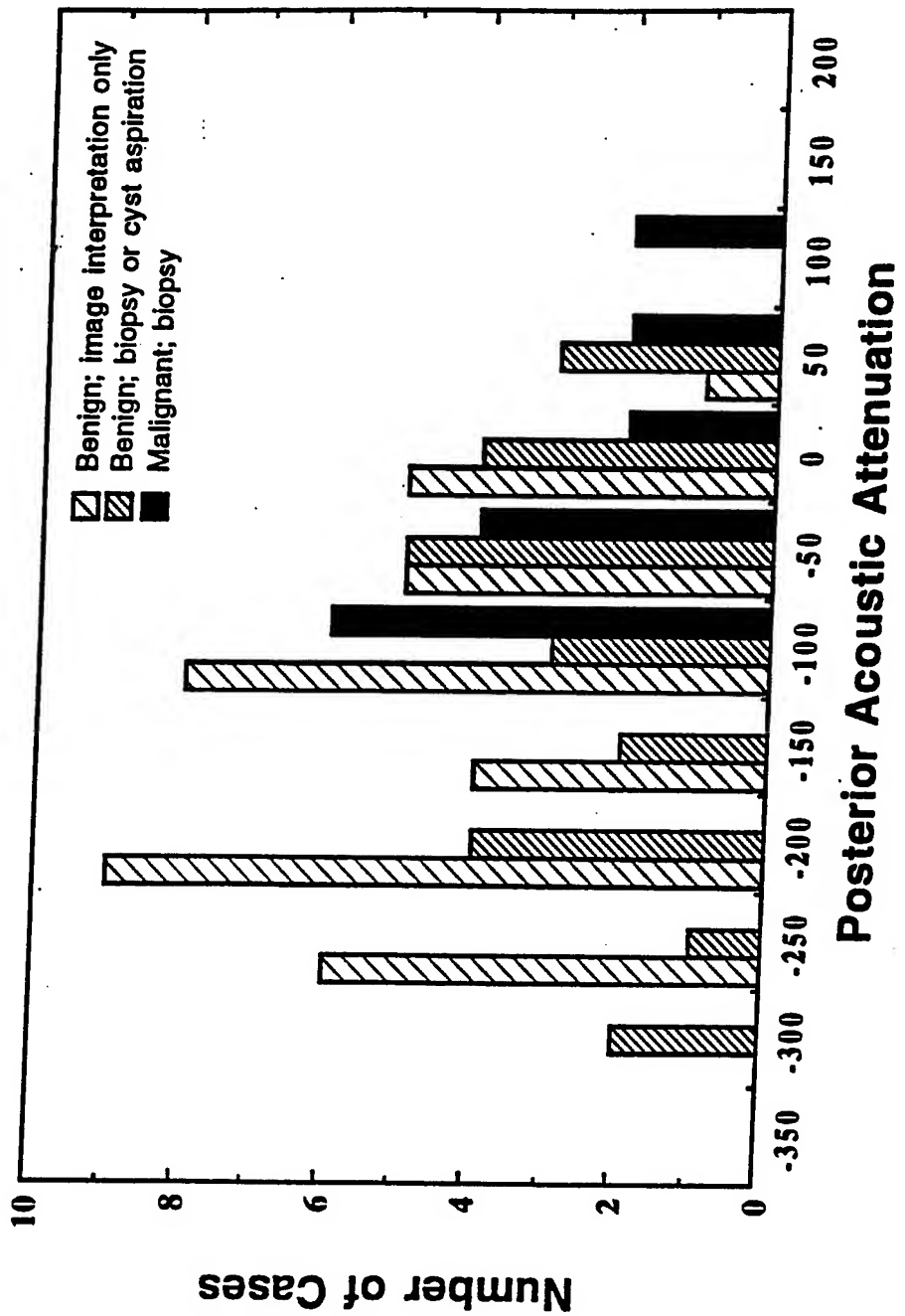

True-Positive Fraction
False-Positive Fraction
Margin Feature
Entire Database
Equivocal Database
0.0
0.5
1.0
0.0
0.5
1.0

Fig. 4a


1.0
0.5
0.0
0.0
0.5
1.0
True-Positive Fraction
False-Positive Fraction
Entire Database
Equivocal Database
Posterior Acoustic Attenuation


Benign; image interpretation only
Benign; biopsy or cyst aspiration
Malignant; biopsy
Number of Cases
0
2
4
6
8
10
12
0.6
0.625
0.65
0.675
0.7
0.725
0.75
0.775
0.8
0.825
0.85
0.875
0.9
Margin Feature


Benign; image interpretation only
Benign; biopsy or cyst aspiration
Malignant; biopsy
Number of Cases
0
2
4
6
8
10
12
0.005
0.006
0.007
0.008
0.009
0.01
0.011
0.012
0.013
0.014
0.015
Coarseness Feature


Benign; image interpretation only
Benign; biopsy or cyst aspiration
Malignant; biopsy
Number of Cases
0
2
4
6
8
10
-350
-300
-250
-200
-150
-100
-50
0
50
100
150
200
Posterior Acoustic Attenuation

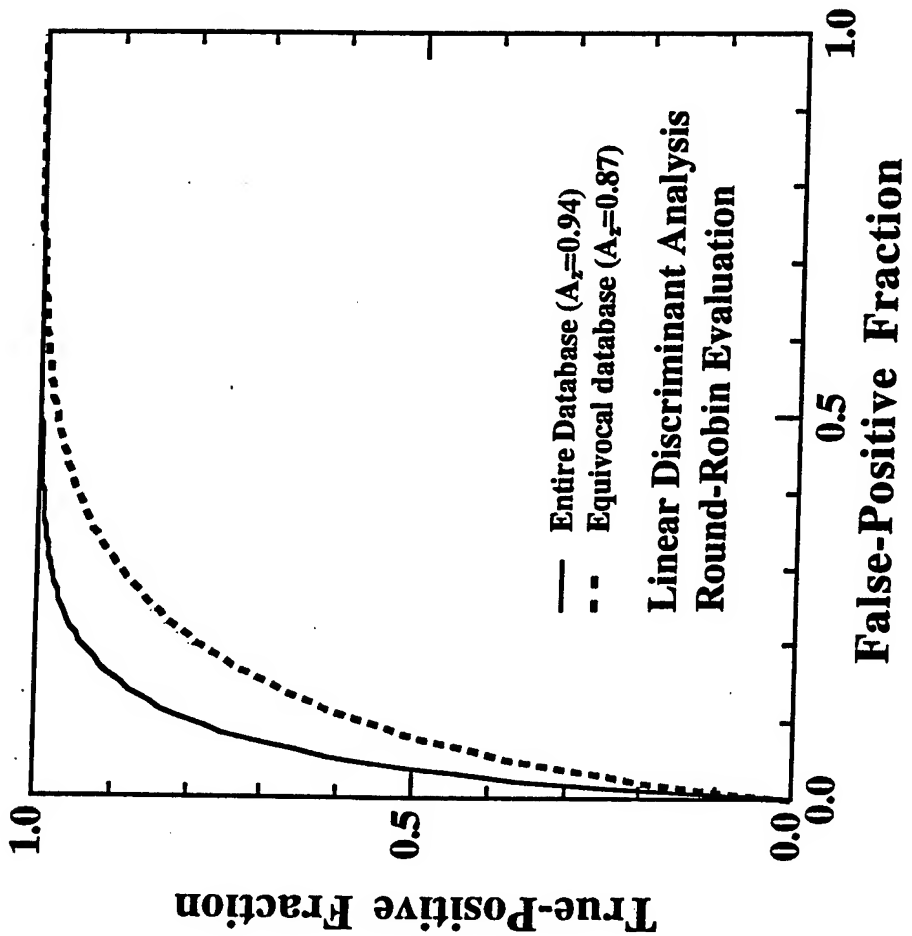

True-Positive Fraction
1.0
0.5
0.0
0.0
0.5
1.0
False-Positive Fraction
Linear Discriminant Analysis
Round-Robin Evaluation
Entire Database (A_z=0.94)
Equivocal database (A_z=0.87)

Fig. 6